Probabilidad y estadística aplicadas a la ingeniería

Segunda edición

Douglas C. Montgomery
Universidad Estatal de Arizona

George C. Runger
Universidad Estatal de Arizona

LIMUSA WILEY

LC	DEWEY	CUTTER
QA276.12	620.015	M 66p

VERSIÓN AUTORIZADA EN ESPAÑOL DE LA OBRA PUBLICADA EN INGLÉS CON EL TÍTULO:
APPLIED STATISTICS AND PROBABILITY FOR ENGINEERS
© JOHN WILEY & SONS, INC., NEW YORK, CHICHESTER, BRISBANE, SINGAPORE, TORONTO AND WEINHEIM.

COLABORADOR EN LA TRADUCCIÓN:
RODOLFO PIÑA GARCÍA

REVISIÓN:
ALMA ROSA GRISELDA ZETINA VÉLEZ
INGENIERA QUÍMICA POR LA FACULTAD DE QUÍMICA DE LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. DOCENTE EN MATEMÁTICAS. JEFA DEL DEPARTAMENTO DE ESTADÍSTICA DE LA UNIDAD DE ADMINISTRACIÓN DEL POSGRADO DGAE-UNAM. PROFESORA EN LA ESCUELA DE CIENCIAS QUÍMICAS DE LA UNIVERSIDAD LA SALLE.

LA PRESENTACIÓN Y DISPOSICIÓN EN CONJUNTO DE

PROBABILIDAD Y ESTADÍSTICA APLICADAS A LA INGENIERÍA

SON PROPIEDAD DEL EDITOR. NINGUNA PARTE DE ESTA OBRA PUEDE SER REPRODUCIDA O TRANSMITIDA, MEDIANTE NINGÚN SISTEMA O MÉTODO, ELECTRÓNICO O MECÁNICO (INCLUYENDO EL FOTOCOPIADO, LA GRABACIÓN O CUALQUIER SISTEMA DE RECUPERACIÓN Y ALMACENAMIENTO DE INFORMACIÓN), SIN CONSENTIMIENTO POR ESCRITO DEL EDITOR.

DERECHOS RESERVADOS:

© 2003, EDITORIAL LIMUSA, S.A. DE C.V.
GRUPO NORIEGA EDITORES
BALDERAS 95, MÉXICO, D.F.
C.P. 06040
8503 8050
01(800) 706 9100
5512 2903
limusa@noriega.com.mx
www.noriega.com.mx

CANIEM NÚM. 121

PRIMERA REIMPRESIÓN
DE LA SEGUNDA EDICIÓN

HECHO EN MÉXICO
ISBN 968-18-5915-4

Prefacio

Toda industria debe esforzarse por mejorar continuamente la calidad de sus productos y servicios si es que pretende competir con éxito tanto en el mercado interno como en el mundial. En este esfuerzo para el mejoramiento de la calidad, los ingenieros y científicos juegan un papel importante pues es a ellos a quienes corresponde el diseño y desarrollo de nuevos productos y de nuevos sistemas y procesos de fabricación, así como el perfeccionamiento de los sistemas existentes. Los recursos estadísticos son una herramienta importante en estas actividades porque proporcionan al ingeniero métodos descriptivos y analíticos para abordar la variabilidad de los datos observados.

Éste es un libro de texto introductorio para un primer curso de probabilidad y estadística aplicadas para estudiantes de las carreras de ingeniería, física y química. Si bien muchos de los métodos que se presentan aquí son fundamentales para el análisis estadístico en otras disciplinas —como negocios y administración, ciencias de la vida y ciencias sociales— este libro está dirigido básicamente a los estudiantes de ingeniería. En nuestra opinión, este enfoque es el que mejor se adapta a las necesidades de los estudiantes de ingeniería porque les permite concentrarse en las diversas aplicaciones de la estadística en este campo. No hemos escatimado esfuerzos para asegurarnos de que todos los ejemplos y los ejercicios son propios de la ingeniería, y en casi todos los casos hemos utilizado ejemplos con datos reales, tomados de material ya publicado, o bien son producto de nuestra experiencia como consultores en este campo.

Consideramos que los estudiantes de todas las ramas de la ingeniería deberían llevar más de un curso de estadística. Desafortunadamente, debido a otras demandas académicas, la mayoría sólo llevará un curso de estadística. Este libro está escrito para que pueda usarse en un solo curso, aunque incluye material suficiente para dos cursos, con la esperanza de que un número mayor de estudiantes de ingeniería conozca las importantes aplicaciones de la estadística en su trabajo cotidiano y opte por un segundo curso. Por otra parte, tenemos la certeza de que este libro también constituirá una útil obra de consulta.

ORGANIZACIÓN DEL LIBRO

Se conserva el nivel matemático un tanto moderado de la primera edición. Hemos comprobado que los estudiantes de ingeniería que han acreditado uno o dos semestres de cálculo no tienen

dificultad para estudiar el libro casi en su totalidad. La intención de éste es ayudar al lector a comprender la metodología y cómo aplicarla, no profundizar en la teoría matemática. En esta edición se ha reorganizado una parte del material y se han reescrito segmentos importantes de varios capítulos.

El capítulo 1 es una introducción al campo de la estadística y a la forma en que los ingenieros utilizan la metodología estadística como parte del proceso de solución de problemas en ingeniería. Se exponen e ilustran métodos simples para resumir y describir datos. En este capítulo también se presentan algunas aplicaciones de la estadística en ingeniería, como la construcción de modelos empíricos, el diseño de experimentos de ingeniería y el monitoreo de procesos de fabricación. En capítulos subsecuentes se abordan estos temas con mayor detalle. En el capítulo 2 se continúa con la descripción de datos y se tratan los diagramas de tallo y hoja, los histogramas, los diagramas de caja, así como varios tipos de diagramas de series de tiempo.

En los capítulos 3, 4, 5 y 6 se cubren los conceptos básicos de probabilidad, variables aleatorias continuas y discretas, valores esperados, distribuciones de probabilidad conjunta e independencia. Se hace un tratamiento razonablemente completo de estos temas, pero se omiten muchos de los detalles matemáticos o de carácter más teórico.

En el capítulo 7 se inicia el tratamiento de la inferencia estadística con la estimación puntual de parámetros. En este capítulo también se introducen los conceptos de muestreo aleatorio, algunas de las propiedades importantes de los estimadores, el método de máxima verosimilitud, las distribuciones muestrales y el teorema del límite central. Se introduce asimismo el *bootstrap* (cargador inicial) como una técnica para encontrar el error estándar de una estimación puntual.

En los capítulos 8 y 9 se exponen la inferencia estadística para una sola muestra y la inferencia estadística para dos muestras, respectivamente. Gran parte del material se ha escrito de nuevo y se le ha dado una nueva organización. Se trata la prueba de hipótesis, así como los intervalos de confianza de medias, varianzas y proporciones, junto con información detallada y ejemplos de los métodos para determinar el tamaño adecuado de la muestra. El propósito aquí es que el estudiante se familiarice con la manera en que se utilizan estas técnicas para resolver problemas de ingeniería del mundo real y que entienda los conceptos que las fundamentan. Se presenta un desarrollo heurístico y lógico de los procedimientos, en vez de recurrir a un desarrollo matemático formal.

En los capítulos 10 y 11 se estudia la regresión lineal simple y múltiple. En la parte sobre regresión múltiple se utiliza el álgebra matricial (capítulo 11), ya que, para ser francos, es la única manera sencilla de entender los conceptos planteados. Las presentaciones aritméticas escalares de la regresión múltiple son confusas y la experiencia nos ha demostrado que los estudiantes de ingeniería cuentan con los conocimientos de álgebra matricial suficientes para entender la presentación de este material.

En los capítulos 12 y 13 se tratan experimentos con un solo factor y con factores múltiples, respectivamente. Se hace hincapié en las nociones de aleatorización, formación de bloques, diseños factoriales, interacciones, análisis de datos gráficos y factoriales fraccionados. En el capítulo 14 se presenta una breve introducción de los métodos y aplicaciones de la estadística no paramétrica, y en el capítulo 15 se introduce al lector al control estadístico de la calidad, enfatizando las cartas de control y los fundamentos del control estadístico de procesos.

Además del conjunto convencional de tablas y diagramas estadísticos, en los apéndices se incluye también material técnico complementario. Este material incluye una introducción a la función generadora de momentos, la técnica del cambio de variable, las permutaciones y los métodos de conteo, el desarrollo de las distribuciones t y F, la estimación de Bayes y el principio

del cociente de verosimilitud. Este material puede resultar de interés para algunos profesores y estudiantes y se presenta como material de referencia.

Cada capítulo contiene una extensa colección de ejercicios, como ejercicios de fin de sección, que hacen énfasis en el material de la misma; ejercicios complementarios al final del capítulo, que cubren la variedad de temas en el mismo, y ejercicios para desarrollar el intelecto, los cuales pueden requerir que el estudiante profundice un poco en el material del libro o que lo aplique de una forma novedosa.

FORMA DE USAR EL LIBRO

Éste es un libro de texto muy flexible, debido a que la idea de los profesores en cuanto a lo que debería incluirse en un primer curso de estadística para ingenieros varía considerablemente, al igual que los antecedentes académicos de los diferentes grupos de estudiantes. Por lo tanto, dudamos en dar demasiados consejos pero explicaremos cómo usamos nosotros el libro.

Consideramos que un primer curso de estadística para ingenieros debe ser ante todo un *curso de estadística aplicada*, no un curso de probabilidad. En un curso de un semestre nosotros cubrimos los capítulos 1 y 2 completos (en tres o cuatro clases), hacemos una revisión general del material sobre probabilidad, haciendo el énfasis principal en la distribución normal (seis a ocho clases), trabajamos la mayor parte de los capítulos 8 y 9 sobre intervalos de confianza y prueba de hipótesis (diez clases), iniciamos la presentación de los modelos de regresión del capítulo 10 (cuatro clases), hacemos una introducción al diseño de experimentos de los capítulos 12 y 13 (seis clases) y presentamos los conceptos básicos del control estadístico de procesos, incluyendo el diagrama de control de Shewhart del capítulo 15 (seis clases). Esto deja entre tres y cuatro periodos para exámenes y repaso. Queremos hacer énfasis en que el propósito de este curso es presentar a los ingenieros la forma en que puede aplicarse la estadística para resolver problemas de ingeniería del mundo real, no para eliminar a los estudiantes menos dotados para las matemáticas. Éste no es el curso de "estadística matemática elemental" que con demasiada frecuencia se imparte a ingenieros.

Si se cuenta con un segundo semestre, entonces es posible cubrir el libro completo, incluyendo parte del material del apéndice, en caso de ser apropiado para los estudiantes. También es posible asignar y resolver en clase muchos de los problemas de tarea para reforzar la comprensión de los conceptos. Obviamente, la regresión múltiple y el diseño de más experimentos serían temas importantes en un segundo curso.

USO DE LA COMPUTADORA

En la práctica, los ingenieros utilizan la computadora para aplicar los métodos estadísticos a la solución de problemas. Por lo tanto, se recomienda ampliamente integrar la computadora en el salón de clases. A lo largo del libro se presentan resultados de Minitab y SAS como ejemplos típicos de lo que puede hacerse con el software moderno de estadística. En la enseñanza, no sólo hemos utilizado estos paquetes sino también otros, como EXCEL, STATGRAPHICS, DESIGN-EASE, JMP y SPSS. No quisimos saturar el libro con ejemplos de diversos paquetes diferentes, ya que la forma en que el profesor integre el software en la clase es más importante que el paquete

que utilice. Todos los datos de texto están disponibles en forma electrónica. En algunos capítulos se presentan problemas que en nuestra opinión deberían resolverse utilizando software de computadora. Estos problemas están identificados con un icono especial al margen.

Ya en el aula, a casi todas las clases llevamos una PC notebook con pantalla de cristal líquido y mostramos cómo se implementa la técnica en la computadora tan pronto como ésta es explicada en clase. Existen versiones para el estudiante de muchos paquetes de software de estadística y los estudiantes pueden adquirir su propia copia o bien usar los productos disponibles en las redes de área local para PC. Hemos comprendido que con esto se consiguen mejoras sensibles tanto para el avance del curso como para que el estudiante comprenda el contenido de este curso.

USO DE LA WEB

En www.wiley.com/college/montASPE/ pueden encontrarse recursos adicionales para estudiantes y profesores.

RECONOCIMIENTOS

Queremos expresar nuestro agradecimiento a las diferentes organizaciones e individuos que han contribuido a la preparación de este libro. Muchos profesores que usaron la primera edición nos hicieron excelentes sugerencias que hemos incorporado en esta edición revisada. También estamos en deuda con el Dr. Smiley Cheng por el permiso para adaptar muchas de las tablas estadísticas de su excelente libro (en colaboración con el Dr. James Fu), *Statistical Tables for Classroom and Exam Room*. John Wiley and Sons, Prentice-Hall, el Institute of Mathematical Statistics y los editores de *Biometrics* nos permitieron usar material protegido por derechos de autor, por lo que expresamos nuestro agradecimiento.

Douglas C. Montgomery
George C. Runger

Contenido

CAPÍTULO 1 El papel de la estadística en la ingeniería 1

1-1 El método de la ingeniería y el pensamiento estadístico 1
1-1.1 La ingeniería y la solución de problemas 1
1-1.2 El pensamiento estadístico 9
1-2 Recolección de datos de ingeniería 11
1-3 Modelos mecanicista y empírico 13
1-4 Diseño de investigaciones experimentales 16
1-5 Observación de procesos en el tiempo 20
Ejercicios complementarios 24
Ejercicios para desarrollar el intelecto 25

CAPÍTULO 2 Resumen y presentación de datos 26

2-1 La importancia del resumen y la presentación de datos 26
2-2 Diagramas de tallo y hoja 27
2-3 Distribuciones de frecuencia e histogramas 32
2-4 Gráficas de caja 37
2-5 Gráficas de series de tiempo 40
Ejercicios complementarios 44
Ejercicios para desarrollar el intelecto 48

CAPÍTULO 3 Probabilidad 49

3-1 Espacios muestrales y eventos 49
3-1.1 Introducción 49
3-1.2 Espacios muestrales 52
3-1.3 Eventos 54
3-2 Interpretación de la probabilidad 62
3-2.1 Introducción 62
3-2.2 Axiomas de probabilidad 66
3-3 Reglas de adición 68
3-4 Probabilidad condicional 74
3-5 Reglas de multiplicación y de probabilidad total 79
3-5.1 Regla de multiplicación 79
3-5.2 Regla de probabilidad total 79
3-6 Independencia 83
3-7 Teorema de Bayes 89
3-8 Variables aleatorias 91
Ejercicios complementarios 93
Ejercicios para desarrollar el intelecto 96

CAPÍTULO 4 Variables aleatorias discretas y distribuciones de probabilidad 97

4-1 Variables aleatorias discretas 97
4-2 Distribuciones de probabilidad y funciones de masa de probabilidad 99
4-3 Funciones de distribución acumulada 103
4-4 Media y varianza de una variable aleatoria discreta 106
4-5 Distribución discreta uniforme 110
4-6 Distribución binomial 112
4-7 Distribuciones geométrica y binomial negativa 121
4-7.1 Distribución geométrica 121
4-7.2 Distribución binomial negativa 124
4-8 Distribución hipergeométrica 129
4-9 Distribución de Poisson 135
Ejercicios complementarios 141
Ejercicios para desarrollar el intelecto 143

CAPÍTULO 5 Variables aleatorias continuas y distribuciones de probabilidad 144

5-1 Variables aleatorias continuas 144
5-2 Distribuciones de probabilidad y funciones de densidad de probabilidad 145
5-3 Funciones de distribución acumulada 149
5-4 Media y varianza de una variable aleatoria continua 153
5-5 Distribución continua uniforme 155
5-6 Distribución normal 157
5-7 Gráficas de probabilidad 170
5-8 Aproximación normal a la distribución binomial y de Poisson 174
5-9 Distribución exponencial 179
5-10 Distribuciones de Erlang y gamma 187
5-10.1 Distribución de Erlang 187
5-10.2 Distribución gamma 189
5-11 Distribución de Weibull 193
Ejercicios complementarios 195
Ejercicios para desarrollar el intelecto 197

CAPÍTULO 6 Distribución de probabilidad conjunta 199

6-1 Dos variables aleatorias discretas 200
6-1.1 Distribuciones de probabilidad conjunta 200
6-1.2 Distribuciones de probabilidad marginal 202
6-1.3 Distribuciones de probabilidad condicional 205
6-1.4 Independencia 207
6-2 Variables aleatorias discretas múltiples 211
6-2.1 Distribuciones de probabilidad conjunta 211
6-2.2 Distribución de probabilidad multinomial 214
6-3 Dos variables aleatorias continuas 218
6-3.1 Distribuciones de probabilidad conjunta 218
6-3.2 Distribuciones de probabilidad marginal 222
6-3.3 Distribuciones de probabilidad condicional 224
6-3.4 Independencia 228
6-4 Variables aleatorias continuas múltiples 231
6-5 Covarianza y correlación 237
6-6 Distribución normal de dos variables 245
6-7 Combinaciones lineales de variables aleatorias 249
6-8 Desigualdad de Chebyshev 254
Ejercicios complementarios 256
Ejercicios para desarrollar el intelecto 259

CAPÍTULO 7 Estimación de parámetros 260

7-1 Inferencia estadística 260
7-2 Muestreo aleatorio 261
7-3 Propiedades de los estimadores 265
7-3.1 Estimadores insesgados 265
7-3.2 Varianza de un estimador puntual 267
7-3.3 Error estándar: reporte de una estimación puntual 268
7-3.4 Estimación *bootstrap* (cargador inicial) del error estándar 270
7-3.5 Error cuadrado medio de un estimador 271
7-4 Método de máxima verosimilitud 274
7-5 Distribuciones de muestreo 281
7-6 Distribuciones de muestreo de medias 282
7-7 Introducción a los intervalos de confianza 289
Ejercicios complementarios 291
Ejercicios para desarrollar el intelecto 292

CAPÍTULO 8 Inferencia estadística para una sola muestra 294

8-1 Prueba de hipótesis 295
8-1.1 Hipótesis estadísticas 295
8-1.2 Pruebas de hipótesis estadísticas 296
8-1.3 Hipótesis de una y dos colas 304
8-1.4 Procedimiento general de las pruebas de hipótesis 307

8-2 Inferencia sobre la media de una población, varianza conocida 309
8-2.1 Pruebas de hipótesis sobre la media 309
8-2.2 Valores P en las pruebas de hipótesis 312
8-2.3 Error tipo II y elección del tamaño de la muestra 313
8-2.4 Prueba para muestras grandes 318
8-2.5 Algunos comentarios prácticos sobre las pruebas de hipótesis 318
8-2.6 Intervalo de confianza para la media 319
8-2.7 Método general para deducir un intervalo de confianza 324
8-2.8 Intervalos de confianza *bootstrap* (cargador inicial) 325
8-3 Inferencia sobre la media de una población, varianza desconocida 328
8-3.1 Pruebas de hipótesis sobre la media 329
8-3.2 Valor P para una prueba t 334
8-3.3 Solución de computadora 334
8-3.4 Elección del tamaño de la muestra 335
8-3.5 Intervalo de confianza para la media 336
8-4 Inferencia sobre la varianza de una población normal 340
8-4.1 Pruebas de hipótesis sobre la varianza de una población normal 340
8-4.2 Error β y elección del tamaño de la muestra 344
8-4.3 Intervalo de confianza sobre la varianza de una población normal 345
8-5 Inferencia sobre una proporción de una población 347
8-5.1 Pruebas de hipótesis sobre una proporción binomial 348
8-5.2 Error tipo II y elección del tamaño de la muestra 349
8-5.3 Intervalo de confianza para una proporción binomial 351
8-6 Tabla de resumen para procedimientos de inferencia para una sola muestra 356
8-7 Prueba de la bondad del ajuste 356
8-8 Pruebas de tablas de contingencia 361
Ejercicios complementarios 365
Ejercicios para desarrollar el intelecto 371

CAPÍTULO 9 Inferencia estadística para dos muestras 375

9-1 Introducción 376
9-2 Inferencia para la diferencia en medias, varianzas conocidas 376
9-2.1 Pruebas de hipótesis para la diferencia en medias, varianzas conocidas 377
9-2.2 Elección del tamaño de la muestra 379
9-2.3 Identificación de causa y efecto 382
9-2.4 Intervalo de confianza para una diferencia en medias, varianzas conocidas 382
9-3 Inferencia para la diferencia en medias de dos distribuciones normales, varianzas desconocidas 388
9-3.1 Pruebas de hipótesis para la diferencia en medias, varianzas desconocidas 388
9-3.2 Elección del tamaño de la muestra 394
9-3.3 Intervalo de confianza para la diferencia en medias 395
9-3.4 Solución de computadora 397
9-4 Prueba t pareada 402
9-5 Inferencia sobre las varianzas de dos poblaciones normales 410
9-5.1 Pruebas de hipótesis sobre el cociente de dos varianzas 410
9-5.2 Error β y elección del tamaño de la muestra 414
9-5.3 Intervalo de confianza sobre el cociente de dos varianzas 415
9-6 Inferencia sobre proporciones de dos poblaciones 417
9-6.1 Prueba de una muestra grande para $H_0: p_1 = p_2$ 417
9-6.2 Error β y elección del tamaño de la muestra 419
9-6.3 Intervalo de confianza para $p_1 - p_2$ 421
9-7 Tabla de resumen para procedimientos de inferencia para dos muestras 423
Ejercicios complementarios 423
Ejercicios para desarrollar el intelecto 429

CAPÍTULO 10 Regresión lineal simple y correlación 430

10-1 Modelos empíricos 431
10-2 Regresión lineal simple 434
10-3 Propiedades de los estimadores de mínimos cuadrados y la estimación de σ^2 444
10-4 Abusos comunes de la regresión 446
10-5 Pruebas de hipótesis en la regresión lineal simple 447
10-5.1 Uso de las pruebas t 447
10-5.2 Enfoque del análisis de varianza para probar la significación de una regresión 450
10-6 Intervalos de confianza 454
10-6.1 Intervalos de confianza para la pendiente y la ordenada al origen 454
10-6.2 Intervalo de confianza para la respuesta media 455
10-7 Predicción de nuevas observaciones 458
10-8 Evaluando la adecuación del modelo de regresión 461
10-8.1 Análisis de residuales 461
10-8.2 Coeficiente de determinación (R^2) 464
10-8.3 Prueba de falta de ajuste 465
10-9 Transformaciones a una línea recta 471
10-10 Correlación 472
Ejercicios complementarios 478
Ejercicios para desarrollar el intelecto 481

CAPÍTULO 11 Regresión lineal múltiple 483

11-1 Modelo de regresión lineal múltiple 484
11-2 Estimación de los parámetros por mínimos cuadrados 488
11-3 Enfoque matricial de la regresión lineal múltiple 491
11-4 Propiedades de los estimadores de mínimos cuadrados y estimación de σ^2 504
11-5 Pruebas de hipótesis en la regresión lineal múltiple 506
11-5.1 Prueba para la significación de una regresión 506
11-5.2 Pruebas para los coeficientes de regresión individuales y para subconjuntos de coeficientes 509
11-6 Intervalos de confianza en la regresión lineal múltiple 515
11-6.1 Intervalos de confianza para los coeficientes individuales de una regresión 515
11-6.2 Intervalo de confianza para la respuesta media 516
11-7 Predicción de nuevas observaciones 518
11-8 Medidas de la adecuación del modelo 521
11-8.1 Coeficiente de determinación múltiple (R^2) 521
11-8.2 Análisis de residuales 522
11-8.3 Observaciones influyentes 526
11-9 Modelos de regresión polinomial 530
11-10 Variables indicadoras 533
11-11 Selección de variables en regresiones múltiples 539
11-11.1 Problema de la construcción del modelo 539
11-11.2 Procedimientos computacionales para la selección de variables 539
11-11.3 Salida de computadora de una regresión por pasos 549
11-12 Colinealidad múltiple 554
Ejercicios complementarios 556
Ejercicios para desarrollar el intelecto 560

CAPÍTULO 12 Diseño y análisis de experimentos con un solo factor: el análisis de varianza 561

12-1 Estrategia de experimentación 562
12-2 Experimento con un solo factor completamente aleatorizado 563
12-2.1 Un ejemplo 563
12-2.2 El análisis de varianza 564
12-2.3 Salida de computadora 573
12-2.4 Análisis de residuales y verificación del modelo 576

12-3 Pruebas sobre medias de tratamientos individuales 581
12-3.1 Comparación gráfica de medias 581
12-3.2 Contrastes ortogonales 583
12-3.3 Método de la diferencia significativa mínima de Fisher 586
12-4 Modelo de efectos aleatorios 589
12-5 Diseño de bloques completos aleatorizados 595
12-5.1 Diseño y análisis estadístico 595
12-5.2 Pruebas sobre medias de tratamientos individuales 602
12-5.3 Análisis de residuales y verificación del modelo 604
12-5.4 Diseño de bloques completos aleatorizados con factores aleatorios 605
12-6 Determinación del tamaño de la muestra en experimentos con un solo factor 609
12-6.1 Caso con efectos fijos 609
12-6.2 Caso con efectos aleatorios 610
Ejercicios complementarios 612
Ejercicios para desarrollar el intelecto 615

CAPÍTULO 13 Diseño de experimentos con varios factores 617

13-1 Introducción 618
13-2 Algunas aplicaciones de las técnicas de diseño experimental 618
13-3 Experimentos factoriales 621
13-4 Experimentos factoriales con dos factores 627
13-4.1 Análisis estadístico del modelo con efectos fijos 627
13-4.2 Verificación de la adecuación del modelo 635
13-4.3 Salida de computadora 635
13-4.4 Una observación por celda 636
13-4.5 Factores aleatorios 638
13-5 Experimentos factoriales generales 641
13-6 Diseño factorial 2^k 647
13-6.1 Diseño 2^2 648
13-6.2 Diseño 2^k para $k \geq 3$ factores 655
13-6.3 Una sola réplica del diseño 2^k 666
13-6.4 Adición de los puntos centrales a un diseño 2^k 671
13-7 Formación de bloques y confusión en el diseño 2^k 677
13-8 Réplica fraccionada del diseño 2^k 684
13-8.1 Media fracción del diseño 2^k 684
13-8.2 Fracciones menores: el factorial fraccionado 2^{k-p} 693
13-9 Métodos y diseño de superficies de respuesta 705
13-9.1 Método del ascenso más pronunciado 707
13-9.2 Análisis de una superficie de respuesta de segundo orden 711
Ejercicios complementarios 720
Ejercicios para desarrollar el intelecto 725

CAPÍTULO 14 Estadística no paramétrica 726

14-1 Introducción 726
14-2 Prueba de signos 727
14-2.1 Descripción de la prueba 727
14-2.2 Prueba de signos para muestras pareadas 732
14-2.3 Error tipo II para la prueba de signos 734
14-2.4 Comparación con la prueba t 735
14-3 Prueba de rangos con signo de Wilcoxon 738
14-3.1 Descripción de la prueba 738
14-3.2 Aproximación para muestras grandes 740
14-3.3 Observaciones pareadas 740
14-3.4 Comparación con la prueba t 742
14-4 Prueba de la suma de rangos de Wilcoxon 743
14-4.1 Descripción de la prueba 743
14-4.2 Aproximación con muestras grandes 746
14-4.3 Comparación con la prueba t 746

14-5 Métodos no paramétricos en el análisis de varianza 747
14-5.1 Prueba de Kruskal-Wallis 747
14-5.2 Transformación de rangos 750
Ejercicios complementarios 752
Ejercicios para desarrollar el intelecto 753

CAPÍTULO 15 Control estadístico de la calidad 754

15-1 La estadística y el mejoramiento de la calidad 755
15-2 Control estadístico de la calidad 756
15-3 Control estadístico de procesos 756
15-4 Introducción a las cartas de control 757
15-4.1 Principios básicos 757
15-4.2 Diseño de una carta de control 762
15-4.3 Subgrupos racionales 763
15-4.4 Análisis de patrones en cartas de control 765
15-5 Cartas de control $\overline{X}$ y R 768
15-6 Cartas de control para mediciones individuales 776
15-7 Capacidad de un proceso 781
15-8 Cartas de control de atributos 787
15-8.1 Carta P (carta de control de proporciones) 787
15-8.2 Carta U (carta de control de defectos por unidad) 790
15-9 Desempeño de las cartas de control 794
15-10 Carta de control de suma acumulada 797
15-11 Otras herramientas del control estadístico de procesos para resolver problemas 806
15-12 Implementación del control estadístico de procesos 809
Ejercicios complementarios 811
Ejercicios para desarrollar el intelecto 815

APÉNDICES 819

A Tablas y cartas estadísticas A-1
B Material técnico complementario B-1
I Técnicas de conteo B-1
II Función generadora de momentos B-7
III Funciones de variables aleatorias B-15
IV Desarrollo de las distribuciones t y F B-24
V Enfoque bayesiano de la estimación B-28
VI Pruebas del cociente de verosimilitud B-32
VII Factores aleatorios en experimentos factoriales B-35
C Bibliografía C-1
D Respuestas de ejercicios seleccionados D-1

ÍNDICE I-1

1 El papel de la estadística en la ingeniería

PLAN GENERAL DEL CAPÍTULO

1-1 EL MÉTODO DE LA INGENIERÍA Y EL PENSAMIENTO ESTADÍSTICO

1-1.1 La ingeniería y la solución de problemas

1-1.2 El pensamiento estadístico

1-2 RECOLECCIÓN DE DATOS DE INGENIERÍA

1-3 MODELOS MECANICISTA Y EMPÍRICO

1-4 DISEÑO DE INVESTIGACIONES EXPERIMENTALES

1-5 OBSERVACIÓN DE PROCESOS EN EL TIEMPO

1-1 EL MÉTODO DE LA INGENIERÍA Y EL PENSAMIENTO ESTADÍSTICO

1-1.1 La ingeniería y la solución de problemas

Un ingeniero es alguien que resuelve problemas de interés para la sociedad mediante la aplicación eficiente de principios científicos. Los ingenieros llevan a cabo esta tarea perfeccionando un producto o un proceso existente o bien diseñando un producto o proceso nuevo que satisfaga las necesidades de los consumidores. El **método de la ingeniería** o **científico** es el enfoque aplicado para formular y resolver estos problemas. Los pasos del método de la ingeniería son los siguientes:

1. Desarrollar una descripción clara y concisa del problema.
2. Identificar, al menos de manera tentativa, los factores importantes que afectan el problema o que pueden jugar un papel en su solución.

3. Proponer un modelo para el problema, utilizando los conocimientos científicos o de la ingeniería del fenómeno bajo estudio. Consignar todas las limitaciones y/o supuestos del modelo.
4. Realizar los experimentos apropiados y recolectar datos para probar o validar el modelo tentativo o las conclusiones planteadas en los pasos 2 y 3.
5. Refinar el modelo con base en los datos observados.
6. Manipular el modelo para contribuir a desarrollar una solución del problema.
7. Realizar un experimento apropiado para confirmar que la solución propuesta del problema es efectiva a la vez que eficiente.
8. Sacar conclusiones o hacer recomendaciones con base en la solución del problema.

En la figura 1-1 se muestran los pasos del método de la ingeniería. Obsérvese que el método de la ingeniería implica una estrecha interacción entre el problema, los factores que pueden influir en la solución, un modelo del fenómeno y la experimentación para verificar la adecuación del modelo y la solución propuesta del problema. Los pasos 2 al 4 de la figura 1-1 están encerrados en un recuadro, indicando que pueden requerirse varios ciclos o iteraciones de los mismos para obtener la solución final. Por consiguiente, los ingenieros deben conocer una manera eficiente para planear experimentos, recolectar datos, analizar e interpretar datos, y entender cómo se relacionan los datos observados con el modelo que han propuesto para el problema bajo estudio.

El campo de la **estadística** trata de la recolección, presentación, análisis y uso de datos para tomar decisiones, solucionar problemas y diseñar productos y procesos. Debido a que diversos aspectos del ejercicio de la ingeniería implican trabajar con datos, resulta evidente la importancia de ciertos conocimientos de estadística para cualquier ingeniero. Específicamente, las técnicas estadísticas pueden constituir una poderosa ayuda para diseñar nuevos productos y sistemas, mejorar diseños existentes, así como para diseñar, desarrollar y mejorar procesos de producción.

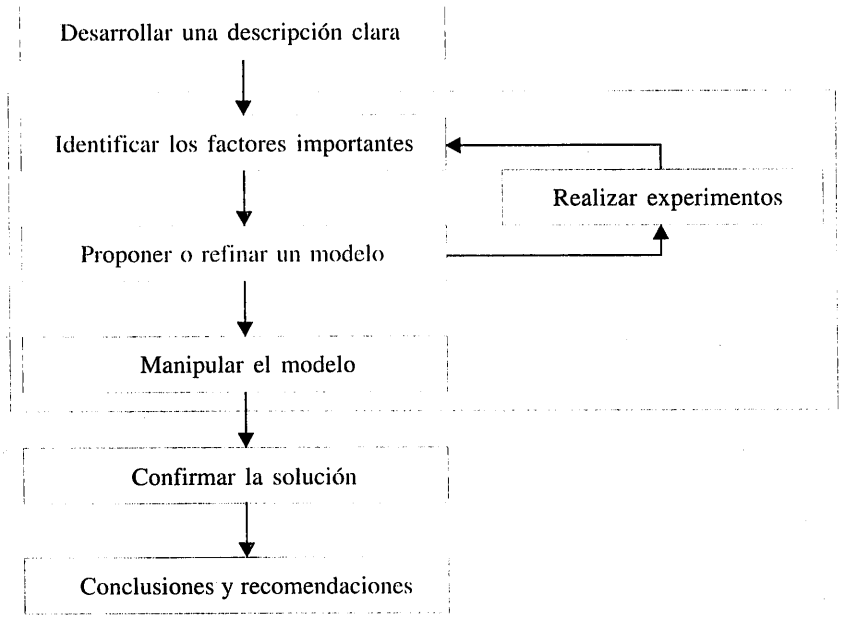


Figura 1-1 El método para resolver problemas.

Los métodos estadísticos se utilizan como ayuda para describir y entender la **variabilidad**. Por variabilidad se entiende que observaciones sucesivas de un sistema o fenómeno no producen exactamente el mismo resultado. Todos nos encontramos con la variabilidad en nuestras vidas cotidianas y el **pensamiento estadístico** puede ofrecernos un recurso conveniente para incorporar esta variabilidad en nuestros procesos de toma de decisiones. Por ejemplo, considérese el rendimiento del tanque de gasolina de un automóvil. ¿Se recorre siempre el mismo kilometraje con cada tanque de combustible? Desde luego que no; de hecho, en ocasiones el kilometraje recorrido varía considerablemente. Esta variabilidad observada en el rendimiento por tanque depende de muchos factores, como las condiciones de manejo que han ocurrido más recientemente (en ciudad o en carretera), los cambios en el estado del vehículo con el tiempo (que podrían incluir factores tales como la presión de las llantas, la compresión del motor o el desgaste de las válvulas), el tipo y octanaje de la gasolina usada, y posiblemente hasta las condiciones meteorológicas que hayan prevalecido recientemente. Estos factores representan **fuentes de variabilidad** potenciales en el sistema. La estadística proporciona un marco para describir esta variabilidad y para saber cuáles de las fuentes de variabilidad son más importantes o cuáles tienen el mayor impacto sobre el rendimiento por tanque de gasolina.

La variabilidad también está presente cuando se abordan problemas de ingeniería. Por ejemplo, supóngase que un ingeniero está diseñando un conector de nylon que se usará en un motor de automóvil. El ingeniero está considerando establecer la especificación de diseño del espesor de la pared en 3/32 de pulgada, pero se encuentra un tanto inseguro acerca del efecto de esta decisión sobre la fuerza de desconexión del conector. Si la fuerza de desconexión es muy débil, el conector puede provocar una avería cuando esté instalado en un motor. Se hacen ocho unidades prototipo y se miden las fuerzas de desconexión, produciendo los siguientes datos (en libras-pie): 12.6, 12.9, 13.4, 12.3, 13.6, 13.5, 12.6, 13.1. Como se anticipaba, la fuerza de desconexión no es la misma para todos los prototipos.

En la figura 1-2 se presenta un **diagrama de puntos** de estos datos. El diagrama de puntos es una gráfica muy útil para ilustrar un número reducido de datos, digamos hasta unas 20 observaciones. Esta gráfica permite ver con facilidad dos características de los datos: la **localización**, o parte de en medio, y la **dispersión** o **variabilidad**. Cuando el número de observaciones es reducido, suele ser difícil identificar algún patrón de variabilidad específico, aun cuando el diagrama de puntos es una forma conveniente para observar cualquier característica inusual de los datos.

También es posible describir numéricamente las características de los datos. Por ejemplo, la localización o tendencia central de los datos puede caracterizarse con el promedio aritmético ordinario o media. Debido a que lo más común es considerar los datos como una muestra, nos referiremos a la media aritmética como la **media muestral**.

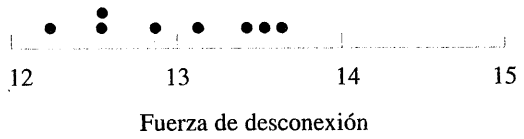


Figura 1-2 Diagrama de puntos de los datos de la fuerza de desconexión cuando el espesor de las paredes es 3/32 de pulgada.

Definición

Si las n observaciones de una muestra se denotan por $x_1, x_2, \ldots, x_n$, entonces la **media muestral** es

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1-1}$$

EJEMPLO 1-1

La media muestral de la fuerza de desconexión de las ocho observaciones realizadas en los conectores prototipo es

$$x = \frac{x_1 + x_2 + \cdots x_n}{n} = \frac{\sum_{i=1}^{8} x_i}{8} = \frac{12.6 + 12.9 + \cdots + 13.1}{8}$$
$$= \frac{104}{8} = 13.0$$

Una interpretación física de la media muestral como medida de localización se muestra en la figura 1-3, que es un diagrama de puntos de los datos de la fuerza de desconexión. Obsérvese que la media muestral $x = 13.0$ puede pensarse como un "punto de equilibrio". Es decir, si cada observación representa 1 libra de masa colocada en el punto correspondiente sobre el eje x, entonces un fulcro localizado en x equilibraría exactamente este sistema de pesos.

La media muestral es el valor promedio de todas las observaciones del conjunto de datos. Por lo general, estos datos son una **muestra** de observaciones que se ha seleccionado de una **población** de observaciones más grande. En este caso, la población podría componerse de todos los conectores que se venderán a los consumidores. En ocasiones existe una población física real, como un lote de pastillas de silicio producidas en una fábrica de semiconductores. También podría considerarse el cálculo del valor promedio de todas las observaciones de una población. A este promedio se le llama la **media poblacional** y se denota por la letra griega μ (mu).

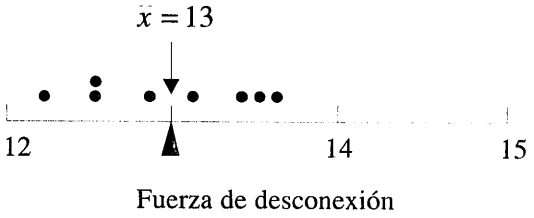


Figura 1-3 La media muestral es el punto de equilibrio de un sistema de pesos.

Cuando hay un número finito de observaciones (digamos N) en la población, entonces la media poblacional es

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \tag{1-2}$$

La media muestral, $\bar{x}$, es una estimación razonable de la media poblacional μ. Por lo tanto, el ingeniero que diseña el conector utilizando un espesor de la pared de 3/32 de pulgada concluiría, con base en los datos, que una estimación de la fuerza de desconexión media es 13.0 libras-pie.

En capítulos posteriores se discutirán los modelos para poblaciones infinitas, lo cual llevará a una definición más general de la media poblacional μ. En muchos problemas de ingeniería importantes es necesario hacer referencia a una media poblacional o tomar decisiones acerca de ella.

Aun cuando la media muestral es útil, no comunica toda la información acerca de una muestra de datos. La variabilidad o dispersión de los datos puede describirse con la **varianza muestral** o la **desviación estándar muestral**.

Definición

Si $x_1, x_2, \ldots, x_n$ es una muestra de n observaciones, entonces la **varianza muestral** es

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \tag{1-3}$$

La **desviación estándar muestral**, s, es la raíz cuadrada positiva de la varianza muestral.

Las unidades de las mediciones de la varianza muestral son las unidades originales de la variable al cuadrado. Por tanto, si x se mide en libras-pie, las unidades de la varianza muestral son (libras-pie)^2. La desviación estándar posee la conveniente propiedad de medir la variabilidad en las unidades originales de la variable de interés, x.

¿Cómo mide la variabilidad la varianza muestral?

Para ver la forma en que la varianza muestral mide la dispersión o variabilidad, remítase a la figura 1-4, donde se muestran las desviaciones $x_i - \bar{x}$ para los datos de la fuerza de desconexión del conector. Entre mayor sea la cantidad de variabilidad en los datos de la fuerza de desconexión, más grandes serán en magnitud absoluta algunas de las desviaciones $x_i - \bar{x}$. Puesto que la suma de las desviaciones $x_i - \bar{x}$ siempre es cero, es necesario usar una medida de variabilidad que cambie las desviaciones negativas a cantidades no negativas. Elevar al cuadrado las desviaciones es el enfoque utilizado en la varianza muestral. Por consiguiente, si s^2 es pequeña, entonces hay relativamente poca variabilidad en los datos, pero si s^2 es grande, la variabilidad es relativamente grande.

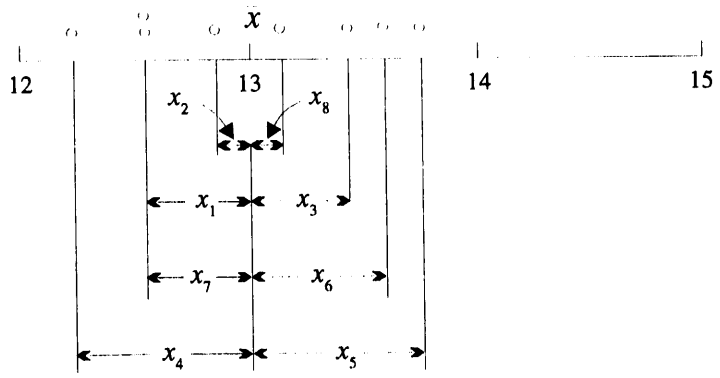


Figura 1-4 Forma en que la varianza muestral mide la variabilidad por medio de las desviaciones $x_i - \bar{x}$.

EJEMPLO 1-2

En la tabla 1-1 se muestran las cantidades necesarias para calcular la varianza muestral y la desviación estándar muestral para los datos de la fuerza de desconexión. Estos datos se grafican en la figura 1-4. El numerador de s^2 es

$$\sum_{i=1}^{8}(x_1 - \bar{x})^2 = 1.60$$

por lo que la varianza muestral es

$$s^2 = \frac{1.60}{8-1} = \frac{1.60}{7} = 0.2286\,(\text{libras} - \text{pie})^2$$

y la desviación estándar muestral es

$$s = \sqrt{0.2286} = 0.48 \text{ libras} - \text{pie}$$

Tabla 1-1 Cálculo de los términos para la varianza muestral y la desviación estándar muestral

i	x_i	$x_i - \bar{x}$	$(x^i - \bar{x})^2$
1	12.6	–0.4	0.16
2	12.9	–0.1	0.01
3	13.4	0.4	0.16
4	12.3	–0.7	0.49
5	13.6	0.6	0.36
6	13.5	0.5	0.25
7	12.6	–0.4	0.16
8	13.1	0.1	0.01
	104.0	0.0	1.60

Cálculo de s^2

Para encontrar el valor de s^2 es necesario hacer el cálculo de $\bar{x}$, de n sustracciones y de n cuadrados y sumar las operaciones. Si las observaciones originales o las desviaciones $x_i - \bar{x}$ no son números enteros, el manejo de las desviaciones $x_i - \bar{x}$ puede ser tedioso y quizá sea necesario trabajar con varios decimales para asegurar la precisión numérica. Una fórmula de cálculo más eficiente para la varianza muestral se obtiene de la siguiente manera:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}\left(x_1^2 + \bar{x}^2 - 2\bar{x}x_i\right)}{n-1} = \frac{\sum_{i=1}^{n}x_i^2 + n\bar{x}^2 - 2\bar{x}\sum_{i=1}^{n}x_i}{n-1}$$

y puesto que $x = (1/n)\sum_{i=1}^{n} x_i$, esta última ecuación se reduce a

$$s^2 = \frac{\sum_{i=1}^{n}x_i^2 - \frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}}{n-1} \tag{1-4}$$

Nótese que la ecuación (1-4) requiere elevar al cuadrado cada x_i individual, después elevar al cuadrado la suma de las x_i, restar $\left(\sum x_i\right)^2/n$ de $\sum x_i^2$ y, por último, dividir entre $n - 1$. En ocasiones, esta ecuación recibe el nombre de método corto para calcular s^2 (o s).

EJEMPLO 1-3

Se calcularán la varianza y la desviación estándar muestrales utilizando el método corto, ecuación 1-4. De la fórmula se obtiene

$$s^2 = \frac{\sum_{i=1}^{n}x_i^2 - \frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}}{n-1} = \frac{1\,353.6 - \frac{(104)^2}{8}}{7} = \frac{1.60}{7} = 0.2286 \quad \text{(libras-pie)}^2$$

y

$$s = \sqrt{0.2286} = 0.48 \text{ libras-pie}$$

Estos resultados concuerdan exactamente con los que se obtuvieron antes.

Análoga a la varianza muestral s^2, existe una medida de la variabilidad de la población llamada **varianza poblacional**. Se usa la letra griega σ^2 (sigma cuadrada) para denotar la varianza poblacional. La raíz cuadrada positiva de σ^2, o σ, denotará la **desviación estándar poblacional**.

Cuando la población es finita y se compone de N valores, la varianza poblacional puede definirse como

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \tag{1-5}$$

Más adelante se dará una definición más general de la varianza σ^2. Se indicó ya que la media muestral puede usarse como una estimación de la media poblacional. De manera similar, la varianza muestral es una estimación de la varianza poblacional.

Obsérvese que el divisor de la varianza muestral es el tamaño de la muestra menos uno $(n - 1)$, en tanto que el de la varianza poblacional es el tamaño de la población N. Si se conociera el valor exacto de la media poblacional μ, entonces la varianza *muestral* podría encontrarse elevando al cuadrado la desviación promedio de las observaciones muestrales alrededor de μ. En la práctica, casi nunca se conoce el valor de μ, por lo que es necesario usar la suma del cuadrado de las desviaciones alrededor del promedio muestral $\bar{x}$. Sin embargo, las observaciones x_i tienden a estar más próximas a su promedio, $\bar{x}$, que a la media poblacional, μ. Por lo tanto, para compensar este hecho se usa $n - 1$ como divisor, en lugar de n. Si se usara n como divisor en la varianza muestral, se obtendría una medida de variabilidad que es, en promedio, consistentemente menor que la verdadera varianza poblacional σ^2.

Otra forma de visualizar lo anterior es considerar la varianza muestral s^2 como una medida basada en $n - 1$ **grados de libertad**. El término *grados de libertad* resulta del hecho de que las n desviaciones $x_1 - \bar{x}$, $x_2 - \bar{x}$, ..., $x_n - \bar{x}$ siempre suman cero, por lo que al especificar los valores de cualesquiera $n - 1$ de estas cantidades automáticamente se determina la única que queda. Esto se ilustró en la tabla 1-1. Por tanto, sólo $n - 1$ de las n desviaciones, $x_i - \bar{x}$, están determinadas libremente.

Además de la varianza muestral y la desviación estándar muestral, el **rango muestral**, la diferencia entre la observación más grande y la más pequeña, es una medida útil de la variabilidad. El rango muestral se define como sigue.

Definición

Si las n observaciones de una muestra se denotan por $x_1, x_2, ..., x_n$, entonces el **rango muestral** es

$$r = \text{máx}(x_i) - \text{mín}(x_i) \tag{1-6}$$

Para los datos de la fuerza de desconexión, el rango muestral es $r = 13.6 - 12.3 = 1.3$. En general, cuando la variabilidad de los datos muestrales aumenta, el rango muestral se incrementa.

El cálculo del rango muestral es sencillo, pero ignora toda la información de los datos muestrales entre el valor más pequeño y el más grande. Por ejemplo, las dos muestras 1, 3, 5, 8, 9 y 1, 5, 5, 5, 9 tienen el mismo rango ($r = 8$). Sin embargo, la desviación estándar de la primera

muestra es $s_1 = 3.35$, en tanto que la de la segunda muestra es $s_2 = 2.83$. La variabilidad es, en efecto, menor en la segunda muestra.

En ocasiones, cuando el tamaño de la muestra es pequeño, digamos $n < 8$ o 10, la pérdida de información asociada con el rango no es demasiado grave. Por ejemplo, el rango tiene un uso generalizado en el control estadístico de calidad, donde los tamaños de las muestras de 4 o 5 observaciones son bastante comunes. En el capítulo 15 se expondrán algunas de estas aplicaciones.

La media, la varianza y la desviación estándar muestrales, así como los diagramas de puntos, son formas simples pero efectivas para resumir los datos. En el capítulo 2 se presentan otros métodos para describir datos.

Con frecuencia, los ingenieros también tienen interés en desarrollar un **modelo** del sistema o proceso que generó los datos. En estos modelos intervienen conceptos de probabilidad que se introducen en el capítulo 3. Se verá que la noción de una distribución de **probabilidad**, como un modelo que describe la variabilidad de un sistema o proceso, es muy importante en el ámbito de la ingeniería. En los capítulos 4 al 6 se explorarán estos conceptos en detalle.

1-1.2 El pensamiento estadístico

La necesidad del pensamiento estadístico surge con frecuencia en la solución de problemas de ingeniería. Considérese al ingeniero que diseña el conector. Después de probar los prototipos, él sabe que una estimación razonable de la fuerza de desconexión media es 13.0 libras-pie. Sin embargo, considera que este valor puede resultar muy bajo para la aplicación proyectada, por lo que decide considerar un diseño alternativo con un espesor de la pared mayor, 1/8 de pulgada. Se construyen ocho prototipos de este diseño y las mediciones observadas de la fuerza de desconexión son 12.9, 13.7, 12.8, 13.9, 14.2, 13.2, 13.5 y 13.1. La media muestral es 13.4 y la desviación estándar muestral es 0.50. Los resultados de las dos muestras se grafican como diagramas de puntos en la figura 1-5. Esta ilustración y los cálculos precedentes dan la impresión de que el aumento del espesor de la pared ha redundado en un incremento en la fuerza de desconexión. Sin embargo, hay algunas preguntas obvias que pueden plantearse. Por ejemplo, ¿cómo sabemos que otra muestra de prototipos no producirá resultados diferentes?, ¿una muestra de ocho prototipos es adecuada para producir resultados confiables? Si los resultados de las pruebas obtenidos hasta el momento se emplean para concluir que al aumentar el espesor de la pared se incrementa la fuerza de desconexión, ¿cuáles son los riesgos asociados con esta decisión? Por ejemplo, ¿es posible que el incremento aparente de la fuerza de desconexión observado en los prototipos con un espesor mayor se deba tan sólo a la variabilidad inherente al sistema y que aumentar el espesor de la pieza (y su costo) no tiene en realidad ningún efecto en la fuerza de desconexión?

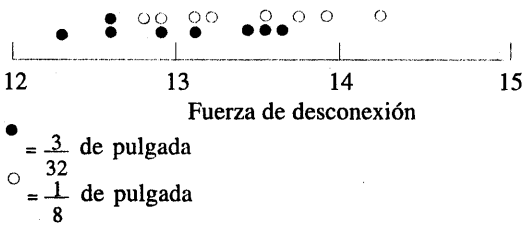


Figura 1-5 Diagrama de puntos de la fuerza de desconexión para los dos espesores de la pared.

En muchos casos se aplican leyes físicas (tales como la ley de Ohm y la ley de los gases ideales) como apoyo para diseñar productos y procesos. Estamos familiarizados con este razonamiento que va de leyes generales a casos específicos. Pero también es importante recurrir a un razonamiento que vaya de una serie de mediciones específicas a casos más generales a fin de responder las preguntas anteriores. Este razonamiento es a partir de una muestra (como los ocho conectores) hacia una población (como los conectores que se venderán a los consumidores). El razonamiento se conoce como **inferencia estadística** (véase la figura 1-6). Históricamente, las mediciones se obtenían de una muestra de habitantes y se generalizaban a una población, terminología que se ha conservado. Desde luego, el razonamiento que lleva de mediciones de algunos objetos a mediciones de todos los objetos pueden resultar en errores (llamados errores de muestreo). Sin embargo, si la selección de la muestra se hace de la manera apropiada, estos riesgos pueden cuantificarse y es posible determinar el tamaño adecuado de la muestra.

En algunos casos, la muestra se selecciona en realidad de la población. La muestra es un subconjunto de la población. Por ejemplo, podría seleccionarse una muestra de tres pastillas de un lote de producción de pastillas de silicio en una fábrica de semiconductores. Con base en los datos de la muestra, quiere concluirse algo acerca del lote. Por ejemplo, no se espera que la media de las mediciones de resistividad de la muestra (x) sea exactamente igual a la media de las mediciones de resistividad del lote (μ). Si embargo, si x tiene un valor alto podría surgir la inquietud de que μ también sea demasiado alta. La inferencia estadística es de x a μ.

En otros casos, la población todavía no existe, pero podría concebirse como futuros duplicados de los objetos de la muestra. Para responder las preguntas anteriores, los ocho conectores prototipo deben ser representativos, en algún sentido, de los que se venderán a los consumidores. En general, los ocho conectores se consideran una muestra de la población compuesta por los conectores que se venderán a los clientes. Desde luego, este análisis requiere alguna noción de estabilidad como supuesto adicional. Por ejemplo, podría suponerse que las fuentes de variabilidad en la fabricación de los prototipos (como la temperatura, la presión y el tiempo de vulcanización) serán las mismas que las de los conectores que se venderán a los consumidores.

Al ejemplo de las pastillas-de-lotes se le llama estudio **enumerativo**. Se usa una muestra para hacer una inferencia sobre la población de la que se ha seleccionado la muestra. Al ejemplo del conector se le llama estudio **analítico**. Se usa una muestra para hacer una inferencia sobre una población futura. Los análisis estadísticos suelen ser los mismos en ambos casos, pero es evidente que un estudio analítico requiere un supuesto de estabilidad (véase la figura 1-7).

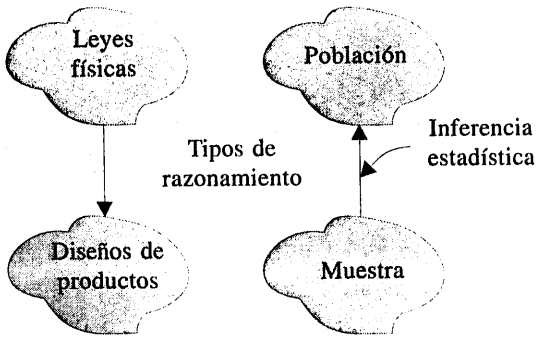


Figura 1-6 La inferencia estadística es uno de los tipos de razonamiento.

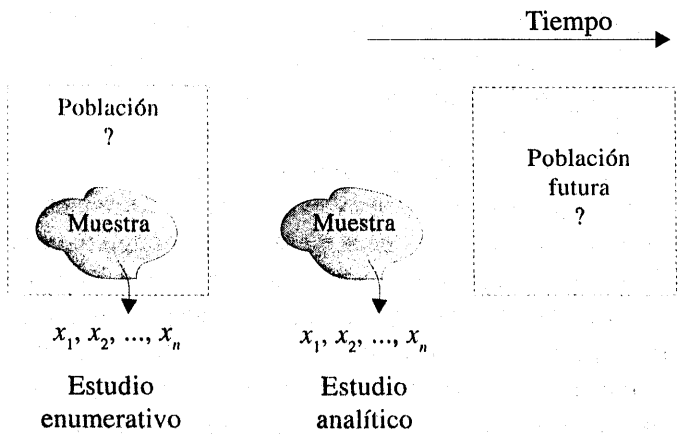


Figura 1-7 Estudio enumerativo y estudio analítico.

1-2 RECOLECCIÓN DE DATOS DE INGENIERÍA

En la sección anterior se ilustraron algunos métodos simples para resumir datos. En el ámbito de la ingeniería, los datos casi siempre son una muestra que se ha seleccionado de una población. En general, estos datos se recolectan en una de dos formas.

La primera forma en que los ingenieros suelen recolectar datos es con un **estudio observacional**. En esta situación, el proceso o sistema que se está estudiando sólo puede ser observado por el ingeniero, y los datos se obtienen conforme se van presentando. Por ejemplo, supóngase que un ingeniero evalúa el desempeño de un proceso de fabricación para hacer componentes de plástico por moldeo por inyección. El ingeniero puede observar el proceso, seleccionar componentes conforme se van fabricando y medir importantes características de interés, tales como el espesor de la pared, la contracción o la resistencia de la pieza. También puede medir y registrar variables del proceso potencialmente importantes, tales como la temperatura del moldeo, el contenido de humedad de la materia prima y el tiempo del ciclo. En un estudio observacional, el ingeniero con frecuencia está interesado en construir un modelo del sistema o proceso. Estos modelos suelen conocerse como modelos empíricos y se introducen e ilustran con mayor detalle en la siguiente sección. Otra forma en que se obtienen datos observacionales es mediante el análisis de los datos históricos sobre el sistema o proceso. Por ejemplo, en la fabricación de semiconductores es muy común llevar extensos registros sobre cada lote de pastillas que se ha producido. Estos registros incluirían datos de prueba de las características físicas y eléctricas de las pastillas, así como las condiciones de procesamiento bajo las cuales se produjo cada lote de pastillas. Si surgen problemas relacionados con un cambio en una característica eléctrica importante, puede estudiarse la historia del proceso en un esfuerzo por determinar el punto en el tiempo en que ocurrió el cambio y tener un conocimiento profundo de las variables del proceso que podrían ser responsables del mismo. Con frecuencia, estos estudios incluyen series de datos muy extensas y requieren una comprensión sólida de los principios estadísticos para que el ingeniero tenga éxito.

La segunda forma en que se obtienen datos de ingeniería es a través de un **experimento diseñado**. En un experimento diseñado, el ingeniero realiza cambios deliberados o intencionados en las variables controlables de un sistema o proceso, observa los datos de los resultados obtenidos del sistema y después hace una inferencia o toma una decisión acerca de las variables

que son responsables de los cambios observados en la eficiencia de los resultados. El ejemplo del conector de nylon de la sección anterior ilustra un experimento diseñado; es decir, se hizo un cambio deliberado en el espesor de la pared del conector con objeto de descubrir si podía obtenerse o no una mayor fuerza de desconexión. Los experimentos diseñados juegan un papel muy importante en los campos de diseño y desarrollo de la ingeniería, así como en el mejoramiento de los procesos de fabricación. En general, cuando el diseño y el desarrollo de productos y procesos se hace utilizando experimentos diseñados, éstos tienen un mejor desempeño, mayor confiabilidad y costos generales más bajos. Los experimentos diseñados también juegan un papel crucial para acortar el tiempo requerido en las actividades de ingeniería de diseño y desarrollo. En la sección 1-4 se ilustrarán varios tipos de experimentos diseñados para el ejemplo del conector.

En la sección 1-1 se introdujeron los conceptos de estudios enumerativos y analíticos. La mayoría de los problemas de ingeniería requieren estudios analíticos. En ambos tipos de estudios pueden obtenerse datos observacionales y datos de experimentos diseñados, pero en la mayoría de los casos incluyen estudios analíticos; es decir, la inferencia o decisión derivada del análisis se refiere al desempeño futuro del sistema o proceso.

La habilidad para visualizar y analizar estadísticamente los datos muestrales permitirá responder preguntas acerca del sistema o proceso bajo estudio. Por ejemplo, considérese el problema de elegir el espesor de la pared del conector de nylon. Un enfoque que podría emplearse para solucionar este problema sería comparar la fuerza de desconexión media para el diseño de 3/32 de pulgada, $\mu_{3/32}$, con la fuerza de desconexión media para el diseño de 1/8 de pulgada, $\mu_{1/8}$, aplicando la técnica de la prueba estadística de hipótesis. En los capítulos 8 y 9 se analiza la prueba de hipótesis y otras técnicas relacionadas. En general, una hipótesis es una proposición acerca de algún aspecto del sistema que es motivo de interés. Por ejemplo, el ingeniero podría interesarse en saber si la fuerza de desconexión media del diseño de 3/32 de pulgada excede la carga máxima típica que se espera encontrar en esta aplicación, digamos 12.75 libras-pie. Por tanto, el interés estaría en probar la hipótesis de que la fuerza media $\mu_{3/32}$ excede 12.75. Éste se conoce como el problema de probar una hipótesis de una sola muestra. En el capítulo 8 se presentan técnicas para este tipo de problema. Alternativamente, el ingeniero podría interesarse en probar la hipótesis de que el incremento del espesor de la pared de 3/32 a 1/8 de pulgada resulta en un incremento de la fuerza de desconexión media. Desde luego, se trata de un estudio analítico; también es un ejemplo del problema de probar una hipótesis con dos muestras. Los problemas de probar hipótesis con dos muestras se exponen en el capítulo 9.

EJERCICIOS DE LAS SECCIONES 1-1 Y 1-2

1-1. Se midió el diámetro interior de ocho anillos forjados para pistones empleados en motores de automóvil. Los datos (en mm) son 74.001, 74.003, 74.015, 74.000, 74.005, 74.002, 74.005 y 74.004. Calcule el promedio muestral y la desviación estándar muestral. Construya un diagrama de puntos y comente los datos.

1-2. En *Applied Life Data Analysis*, Wiley, Wayne Nelson presenta el tiempo de falla de un fluido aislante entre electrodos a 34 kV. Los tiempos, en minutos, son los siguientes: 0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71 y 72.89. Calcule el promedio muestral y la desviación estándar muestral.

1-3. Un número de *Arizona Trend* contiene un suplemento que describe los 12 "mejores" campos de golf del estado. Los yardajes (longitudes) de estos campos son los siguientes: 6 981, 7 099, 6 930, 6 992, 7 518, 7 100, 6 935, 7 518, 7 013, 6 800, 7 041 y 6 890.

Calcule el promedio muestral y la desviación estándar muestral. Construya un diagrama de puntos de los datos.

1-4. Un artículo del *Journal of Structural Engineering* (Vol. 115) describe un experimento para probar la resistencia de tubos circulares con tapones soldados en los extremos. Los primeros datos (en kN) son 96, 96, 102, 102, 102, 104, 104, 108, 126, 126, 128, 128, 140, 156, 160, 160, 164 y 170. Calcule el promedio muestral y la desviación estándar muestral. Construya un diagrama de puntos de los datos.

1-5. Un artículo de *Human Factors* presentó datos sobre el ajuste visual (una función del movimiento del ojo) cuando se identificaba un patrón de manchas en un monitor catódico de alta resolución, CRT. Los datos son los siguientes: 36.45, 67.90, 38.77, 42.18, 26.72, 50.77, 39.30 y 49.71. Calcule el promedio muestral y la desviación estándar muestral. Construya un diagrama de puntos de los datos.

1-6. Los datos siguientes son mediciones directas de la intensidad solar (watts/m^2) en días diferentes en un lugar al sur de España: 562, 869, 708, 775, 775, 704, 809, 856, 655, 806, 878, 909, 918, 558, 768, 870, 918, 940, 946, 661, 820, 898, 935, 952, 957, 693, 835, 905, 939, 955, 960, 498, 653, 730 y 753. Calcule el promedio muestral y la desviación estándar muestral.

1-7. Para los ejercicios del 1-1 al 1-6, comente si los datos resultan de un estudio observacional o de un experimento diseñado.

1-8. Un número de *Aviation Week and Space Technology* reporta que durante la Operación Tormenta del Desierto, los pilotos de los F-117A de la fuerza aérea estadounidense realizaron 1 270 misiones de combate para un total de 6 905 horas. ¿Cuál es la duración media de una misión de un F-117A durante esta operación? ¿Por qué el parámetro calculado es una media poblacional?

1-3 MODELOS MECANICISTA Y EMPÍRICO

Los modelos juegan un papel importante en el análisis de prácticamente todos los problemas de ingeniería. Gran parte de la educación formal de los ingenieros implica el aprendizaje de los modelos relevantes para campos específicos y de las técnicas para aplicar estos modelos en la formulación y solución de problemas. Como un ejemplo simple, supóngase que se mide el flujo de corriente en un alambre delgado de cobre. El modelo para este fenómeno podría ser la ley de Ohm

$$\text{Corriente} = \text{voltaje/resistencia}$$

o bien

$$I = E/R \tag{1-7}$$

A este tipo de modelo se le llama **modelo mecanicista**, porque se construye a partir del conocimiento previo del mecanismo físico básico que relaciona estas variables. Sin embargo, si este proceso de medición se lleva a cabo más de una vez, quizás en tiempos diferentes, o incluso en días diferentes, la corriente observada podría diferir ligeramente debido a pequeños cambios o variaciones en los factores que no están sujetos a un control completo, tales como los cambios en la temperatura ambiente, las fluctuaciones en el funcionamiento del medidor, pequeñas impurezas presentes en diferentes sitios del alambre y las variaciones en la fuente de voltaje. Por consiguiente, un modelo más realista de la corriente observada sería

$$I = E/R + \varepsilon \tag{1-8}$$

donde ε es un término agregado al modelo para tomar en consideración el hecho de que los valores observados del flujo de corriente no se ajustan exactamente al modelo mecanicista. La ε puede visualizarse como un término que incluye los efectos de todas las fuentes de variabilidad no contempladas en el modelo que afectan este sistema.

En ocasiones, los ingenieros trabajan con problemas para los que no existe un modelo mecanicista simple o del todo conocido que explique el fenómeno. Por ejemplo, supóngase que nos interesa el número que expresa el peso molecular medio (M_n) de un polímero. Ahora sabemos que M_n está relacionado con la viscosidad del material (V) y que depende también de la cantidad de catalizador (C) y de la temperatura (T) en el reactor de polimerización cuando se fabrica el material. Una relación entre M_n y estas variables sería, por ejemplo,

$$M_n = f(V, C, T) \tag{1-9}$$

donde no se conoce la *forma* de la función f. Quizás podría desarrollarse un modelo de trabajo a partir de una expansión de primer orden de la serie de Taylor, la cual produciría un modelo de la forma

$$M_n = \beta_0 + \beta_1 V + \beta_2 C + \beta_3 T \tag{1-10}$$

donde las β son parámetros desconocidos. Ahora bien, como en el caso de la ley de Ohm, este modelo no describirá exactamente el fenómeno, por lo que será necesario tomar en consideración las otras fuentes de variabilidad que puedan afectar el peso molecular agregando otro término al modelo, de modo que

$$M_n = \beta_0 + \beta_1 V + \beta_2 C + \beta_3 T + \varepsilon \tag{1-11}$$

es el modelo que se usará para relacionar el peso molecular con las otras tres variables. A un modelo de este tipo se le llama **modelo empírico**; es decir, se aplican en él nuestros conocimientos científicos y de ingeniería del fenómeno, pero no se desarrolla directamente de nuestra comprensión teórica o con base en principios fundamentales del mecanismo subyacente.

Para ilustrar estas ideas con un ejemplo específico, considérense los datos de la tabla 1-2. Esta tabla contiene los datos sobre tres variables recolectados en una fábrica de semiconductores. En esta fábrica, el semiconductor terminado es alambre adherido en una tablilla. Las variables reportadas son la resistencia al desprendimiento (una medida de la cantidad de fuerza requerida para romper la unión), la longitud del alambre y la altura de la matriz. Nos gustaría encontrar un modelo que relacionara la resistencia al desprendimiento con la longitud del alambre y la altura de la matriz. Por desgracia, no hay ningún mecanismo físico que pueda aplicarse fácilmente en este caso, por lo que no parece probable que el enfoque de un modelo mecanicista sea de éxito. Obsérvese que se trata de un ejemplo de un estudio observacional (remítase a la sección 1-2).

La figura 1-8 presenta una gráfica tridimensional de las 25 observaciones de la resistencia al desprendimiento, la longitud del alambre y la altura de la matriz. Del examen de esta gráfica se observa que la resistencia al desprendimiento se incrementa cuando tanto la longitud del alambre como la altura de la matriz se incrementan. Además, parece razonable pensar que el modelo

$$\text{Resistencia al desprendimiento} = \beta_0 + \beta_1 \text{ (longitud del alambre)} + \beta_2 \text{ (altura de la matriz)} + \varepsilon$$

Tabla 1-2 Datos de la resistencia al desprendimiento de la adherencia del alambre

Número de observación	Resistencia al desprendimiento	Longitud del alambre	Altura de la matriz
1	9.95	2	50
2	24.45	8	110
3	31.75	11	120
4	35.00	10	550
5	25.02	8	295
6	16.86	4	200
7	14.38	2	375
8	9.60	2	52
9	24.35	9	100
10	27.50	8	300
11	17.08	4	412
12	37.00	11	400
13	41.95	12	500
14	11.66	2	360
15	21.65	4	205
16	17.89	4	400
17	69.00	20	600
18	10.30	1	585
19	34.93	10	540
20	46.59	15	250
21	44.88	15	290
22	54.12	16	510
23	56.63	17	590
24	22.13	6	100
25	21.15	5	400

sería apropiado como modelo empírico para esta relación. En general, a un modelo empírico de este tipo se le llama **modelo de regresión**. En los capítulos 10 y 11 se muestra la manera de construir estos modelos y de probar su adecuación como funciones de aproximación. Se usará un método para estimar los parámetros de un modelo de regresión, llamado método de mínimos cuadrados, cuyo origen se remonta al trabajo de Karl Gauss. En esencia, en este método se escogen los parámetros del modelo empírico (las β) para minimizar la suma del cuadrado de las

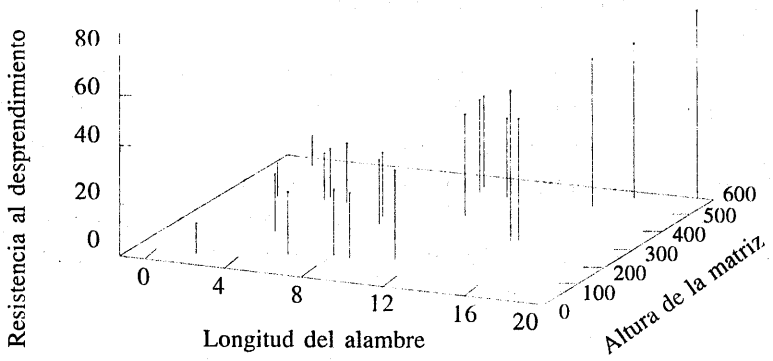


Figura 1-8 Gráfica tridimensional de los datos del alambre y la resistencia al desprendimiento.

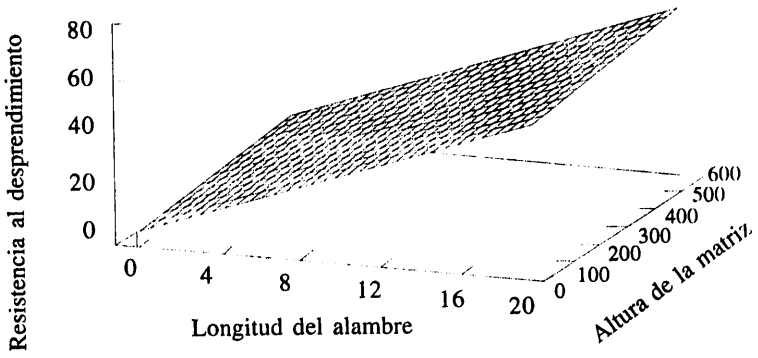


Figura 1-9 Gráfica de los valores predichos de la resistencia al desprendimiento a partir del modelo empírico de la ecuación (1-12).

distancias entre cada punto de los datos y el plano representado por la ecuación del modelo. Resulta que al aplicar esta técnica a los datos de la tabla 1-2 se obtiene

$$\widehat{\text{Resistencia al desprendimiento}} = 2.26 + 2.74\,(\text{longitud del alambre}) + 0.0125\,(\text{altura de la matriz}) \quad (1\text{-}12)$$

donde el "gorro" o acento circunflejo sobre la resistencia al desprendimiento indica que se trata de una cantidad estimada o predicha.

La figura 1-9 es una gráfica de los valores predichos de la resistencia al desprendimiento contra la longitud del alambre y la altura de la matriz obtenidos de la ecuación 1-12. Obsérvese que los valores predichos están sobre un plano arriba del espacio longitud del alambre-altura de la matriz. Por la gráfica de los datos de la figura 1-8, este modelo no parece irrazonable. El modelo empírico de la ecuación 1-12 podría usarse para predecir valores de la resistencia al desprendimiento para varias combinaciones de la longitud del alambre y la altura de la matriz que sean de interés. En esencia, un ingeniero podría usar el modelo empírico exactamente del mismo modo que se usaría un modelo mecanicista.

1-4 DISEÑO DE INVESTIGACIONES EXPERIMENTALES

Gran parte de nuestros conocimientos de ingeniería y ciencias físico-químicas se desarrollan mediante la realización de pruebas o experimentación. Con frecuencia, los ingenieros trabajan en áreas de problemas en las que ninguna teoría científica o de ingeniería tiene una aplicación directa o completa, por lo que la experimentación y la observación de los datos resultantes constituyen la única manera en que puede resolverse el problema. Aun cuando existe una base significativa de teoría científica subyacente en la que podemos apoyarnos para explicar los fenómenos de interés, casi siempre es necesario realizar pruebas o experimentos para confirmar que la teoría en realidad es operativa en la situación o entorno en la que se está aplicando. El pensamiento estadístico y los métodos estadísticos juegan un papel importante en la planeación, realización y análisis de los datos de experimentos de ingeniería.

En la sección 1-1 se trató un breve ejemplo de un ingeniero que investigaba el efecto de aumentar el espesor de la pared de un conector sobre la fuerza de desconexión. Recuérdese que el ingeniero construyó ocho prototipos de cada diseño (3/32 y 1/8 de pulgada), probó cada unidad y calculó el promedio muestral y la desviación estándar muestral de la fuerza de desco-

nexión para cada diseño. Se hizo notar que la prueba de hipótesis estadísticas era un posible marco dentro del cual investigar si un aumento del espesor de la pared en el diseño redundaría en niveles más altos de la fuerza de desconexión media. Esta es una ilustración del uso del pensamiento estadístico como apoyo para el análisis de los datos de un experimento comparativo simple.

El pensamiento estadístico también puede aplicarse a problemas experimentales más complejos. Para ilustrar este punto, reconsidérese el problema del espesor de la pared del conector. Supóngase que cuando el conector se ensambla en la aplicación, primero se sumerge en un adhesivo y luego se vulcaniza el ensamblaje aplicándole calor durante algún periodo. La fuerza de desconexión se mide en el ensamblaje final. El ingeniero sospecha que además del espesor de la pared, el tiempo y la temperatura de vulcanización podrían tener algún efecto sobre el desempeño del conector. Por lo tanto, es necesario un experimento diseñado que permita investigar el efecto de los tres factores sobre la fuerza de desconexión.

Cuando son varios los factores potencialmente importantes, la mejor estrategia de experimentación es diseñar algún tipo de **experimento factorial**. Un experimento factorial es aquel en el que los factores se hacen variar al mismo tiempo. Para ilustrar este punto, supóngase que en el experimento del conector los tiempos de vulcanización de interés son 1 y 24 horas y que los niveles de temperatura son 70°F y 100°F. Ahora bien, dado que los tres factores tienen dos niveles, un experimento factorial constaría de las ocho combinaciones de prueba mostradas en las esquinas del cubo de la figura 1-10. Se realizarían dos ensayos, o **réplicas**, en cada esquina, dando como resultado un experimento factorial de 16 corridas. Los valores observados de la fuerza de desconexión aparecen entre paréntesis en las esquinas del cubo de la figura 1-10. Obsérvese que en este experimento se utilizan ocho prototipos de 3/32 de pulgada y ocho prototipos de 1/8 de pulgada, el mismo número empleado en el estudio comparativo simple de la sección 1-1, pero ahora se están investigando *tres* factores. En general, los experimentos factoriales son la manera más eficiente de estudiar los efectos conjuntos de varios factores.

Es posible sacar interesantes conclusiones tentativas de este experimento. Primero, compárese la fuerza de desconexión promedio de los ocho prototipos de 3/32 de pulgada con la fuerza de desconexión promedio de los ocho prototipos de 1/8 de pulgada (estos son los promedios de las ocho corridas de la cara derecha y la cara izquierda del cubo de la figura 1-10, respectiva-

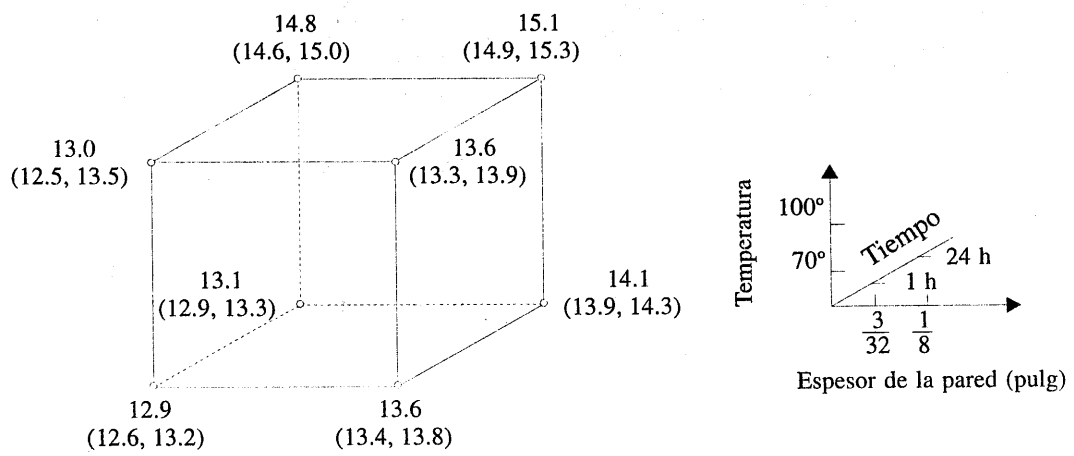


Figura 1-10 El experimento factorial para el problema del espesor de la pared del conector.

mente), o 14.1 – 13.45 = 0.65. Por tanto, al incrementarse el espesor de la pared de 3/32 a 1/8 de pulgada, la fuerza de desconexión promedio se incrementa en 0.65 libras-pie. Siguiente, para medir el efecto del incremento del tiempo de vulcanización, compárese el promedio de las ocho corridas en la cara posterior del cubo (donde el tiempo = 24 h) con el promedio de las ocho corridas de la cara frontal (donde el tiempo = 1 h), o 14.275 – 13.275 = 1. El efecto de aumentar el tiempo de vulcanización de 1 a 24 horas es un incremento de la fuerza de desconexión promedio de 1 libra-pie; es decir, el tiempo de vulcanización aparentemente tiene un efecto que es mayor que el efecto de incrementar el espesor de la pared. El efecto de la temperatura de vulcanización puede evaluarse comparando el promedio de las ocho corridas de la cara superior del cubo (donde la temperatura = 100°F) con el promedio de las ocho corridas de la cara inferior (donde la temperatura = 70°F), o 14.125 – 13.425 = 0.7. Por tanto, el efecto de aumentar la temperatura de vulcanización es un incremento de la fuerza de desconexión promedio de 0.7 libras-pie. Por tanto, si el objetivo del ingeniero es diseñar un conector con una fuerza de desconexión alta, aparentemente existen varias alternativas, como incrementar el espesor de la pared y utilizar las condiciones de vulcanización "estándares" de 1 h y 70°F o bien utilizar el espesor de la pared original de 3/32 de pulgada, pero especificando un tiempo de vulcanización más largo y una temperatura de vulcanización más alta.

Existe una interesante relación entre el tiempo y la temperatura de vulcanización que puede verse examinando la gráfica de la figura 1-11. Esta gráfica se construyó calculando la fuerza de desconexión promedio con las cuatro diferentes combinaciones del tiempo y la temperatura, graficando estos promedios contra el tiempo y, por último, uniendo con líneas rectas los puntos que representan los dos niveles de temperatura. La pendiente de cada una de estas rectas representa el efecto del tiempo de vulcanización sobre la fuerza de desconexión. Obsérvese que las pendientes de estas dos rectas no parecen ser iguales, indicando que el efecto del tiempo de vulcanización es *diferente* para los dos valores de la temperatura de vulcanización. Este es un ejemplo de una **interacción** entre dos factores. La interpretación de esta interacción es bastante directa; si se usa el tiempo de vulcanización estándar (1 h), la temperatura de vulcanización tiene un efecto reducido, pero si se usa el tiempo de vulcanización más largo (24 h), el aumento

Tiempo	Temperatura	Fuerza promedio
1 h	70°F	13.25
1 h	100°F	13.30
24 h	70°F	13.60
24 h	100°F	14.95

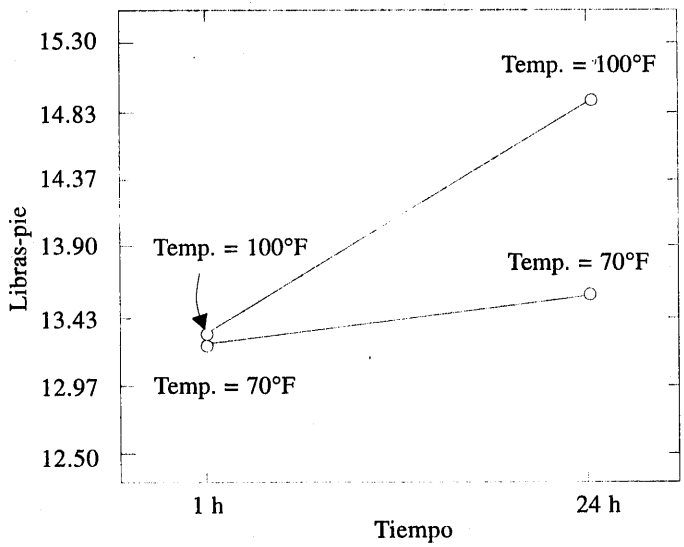


Figura 1-11 La interacción de dos factores entre el tiempo de vulcanización y la temperatura de vulcanización.

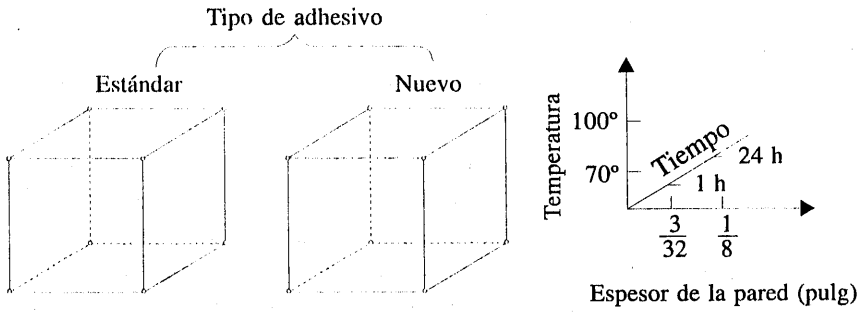


Figura 1-12 Experimento con cuatro factores para el problema del espesor de la pared del conector.

de la temperatura de vulcanización tiene un efecto considerable sobre la fuerza de desconexión promedio. Las interacciones ocurren con frecuencia en sistemas físicos y químicos, y los experimentos factoriales son la única manera de investigar sus efectos. De hecho, si existen interacciones y no se usa la estrategia experimental factorial, pueden obtenerse resultados incorrectos o engañosos.

Es sencillo ampliar la estrategia factorial para un número mayor de factores. Supóngase que el ingeniero quiere considerar un cuarto factor, el tipo de adhesivo. Hay dos tipos: el adhesivo estándar y uno de reciente introducción. En la figura 1-12 se ilustra la forma en que podrían investigarse los cuatro factores —el espesor de la pared, el tiempo de vulcanización, la temperatura de vulcanización y el tipo de adhesivo— en un diseño factorial. Puesto que los cuatro factores siguen teniendo dos niveles, el diseño experimental todavía puede representarse geométricamente como un cubo (se trata en realidad de un *hipercubo*). Obsérvese que, como en cualquier diseño factorial, se prueban todas las combinaciones posibles de los cuatro factores. El experimento requiere 16 ensayos.

En general, si hay k factores presentes y cada uno de ellos tiene dos niveles, un diseño experimental requerirá 2^k corridas. Por ejemplo, con $k = 4$, el diseño 2^4 de la figura 1-12 requiere 16 pruebas. Obviamente, conforme el número de factores aumenta, el número de ensayos requeridos en un experimento factorial se incrementa con rapidez; por ejemplo, ocho factores con dos niveles cada uno requeriría 256 ensayos. Esto pronto deja de ser factible desde el punto de vista del tiempo y otros recursos. Por fortuna, cuando hay entre cuatro y cinco o más factores, por lo general no es necesario probar todas las combinaciones posibles de los niveles de los factores. Un **experimento factorial fraccionado** es una variación del arreglo factorial básico en el que sólo se prueba realmente un subconjunto de las combinaciones de los factores. En la

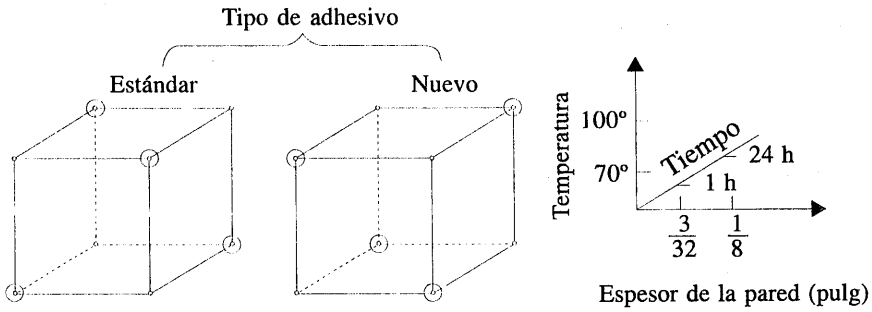


Figura 1-13 Experimento factorial fraccionado para el problema del espesor de la pared del conector.

figura 1-13 se muestra un diseño experimental factorial fraccionado para la versión con cuatro factores del experimento del conector. Las combinaciones de prueba que aparecen encerradas en un círculo en esta figura son las únicas combinaciones de prueba que necesitan correrse. Este diseño experimental sólo requiere ocho corridas en vez de las 16 originales; por consiguiente, se llamaría de **media fracción**. Este es un excelente diseño experimental para estudiar los cuatro factores. Proporcionará información adecuada acerca de los efectos individuales de los cuatro factores, así como cierta información acerca de la manera en que interactúan.

Los experimentos factoriales y factoriales fraccionados son usados ampliamente por ingenieros y científicos en el campo de la investigación y el desarrollo industrial, donde se diseñan y desarrollan nuevas tecnologías, productos y procesos, y donde se mejoran los productos y procesos existentes. Dada la enorme cantidad de trabajos de ingeniería donde interviene la realización de pruebas y la experimentación, es fundamental que todos los ingenieros comprendan los principios básicos de la planeación eficiente y eficaz de experimentos. Estos principios se discuten en el capítulo 12. El capítulo 13 se dedica a los factoriales y los factoriales fraccionados que se han introducido aquí.

1-5 OBSERVACIÓN DE PROCESOS EN EL TIEMPO

Siempre que se recolectan datos en el tiempo es importante graficarlos en esta dimensión. Los fenómenos que podrían afectar el sistema o proceso con frecuencia se hacen más evidentes en una gráfica orientada en el tiempo y el concepto de estabilidad puede juzgarse mejor.

La figura 1-14 es un diagrama de puntos de las lecturas de la concentración tomadas periódicamente en un proceso químico. La gran variación que se observa en el diagrama de puntos indica un posible problema, pero el diagrama no ayuda a explicar la razón de la variación. Debido a que los datos se recolectan en el tiempo, en la figura 1-15 se grafican en el tiempo. En la gráfica es visible un movimiento en el nivel medio del proceso y puede obtenerse una estimación del tiempo en que ocurre el corrimiento.

El gurú de la calidad Edward Deming destacó la importancia de entender la naturaleza de la variación en el tiempo. Realizó un experimento en el que intentaba arrojar unas canicas tan cerca como fuera posible de un blanco sobre una mesa. Utilizó un embudo montado en un aro con pedestal y las canicas se dejaban caer en el embudo (véase la figura 1-16). El embudo se alineó lo más cerca posible con el centro del blanco. Después Deming utilizó dos estrategias diferentes para operar el proceso. 1) Nunca cambiaba de posición el embudo. Simplemente dejaba caer una canica tras otra y registraba la distancia del blanco. 2) Dejaba caer la primera canica y registraba su posición en relación con el blanco. Después movía el embudo una distancia igual pero en dirección opuesta en un intento por compensar el error. Repetía el mismo ajuste después de dejar caer cada canica.

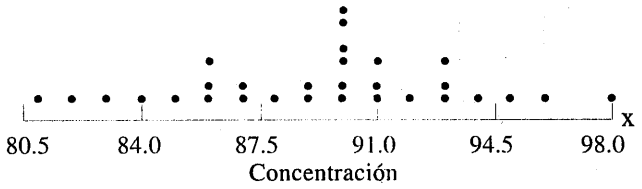


Figura 1-14 El diagrama de puntos ilustra la variación, pero no identifica el problema.

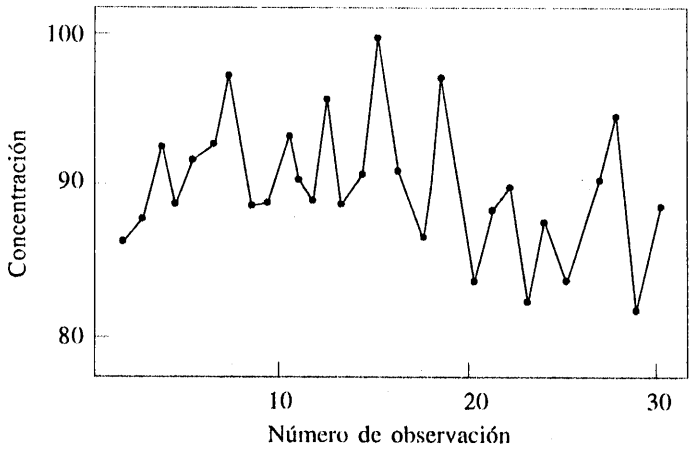


Figura 1-15 Una gráfica de la concentración contra el tiempo.

Después de llevar a cabo las dos estrategias, Deming observó que la varianza de la distancia del blanco para la estrategia 2 era aproximadamente dos veces mayor que para la estrategia 1. Los ajustes del embudo incrementaron las desviaciones del blanco. La explicación es que el error (la desviación de la posición de la canica respecto del blanco) para una canica no proporciona información sobre el error que ocurrirá en la canica siguiente. Por consiguiente, los ajustes del embudo no disminuyen los errores futuros. Más bien, tienden a alejar más el embudo del blanco.

Este interesante experimento indica que los ajustes de un proceso basado en perturbaciones aleatorias en realidad pueden *incrementar* la variación del proceso. Este hecho se conoce como **sobrecontrol** o **corrupción**. Los ajustes sólo deberán aplicarse para compensar un cambio no aleatorio en el proceso; entonces pueden ayudar. Es posible usar una simulación de computadora para poner de manifiesto las enseñanzas del experimento del embudo. En la figura 1-17*a* se muestra una gráfica en el tiempo de 50 mediciones (denotadas por y) de un proceso en el que únicamente están presentes perturbaciones aleatorias. El valor al que se quiere llegar en el proceso es 10 unidades. En la figura 1-17*b* se muestran los mismos datos después de aplicar ajustes a la media del proceso en un intento por producir datos más próximos al valor que quiere alcanzarse. Cada ajuste es igual y en la dirección opuesta a la desviación mostrada por la medición anterior respecto del valor al que quiere llegarse. Por ejemplo, cuando la medición es 11 (una unidad arriba del valor que quiere alcanzarse), la media se reduce una unidad antes de generar la medición siguiente. El sobrecontrol ha incrementado las desviaciones respecto del valor que quiere alcanzarse.

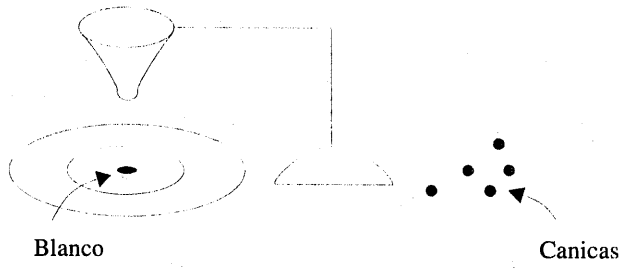


Figura 1-16 Experimento del embudo de Deming.

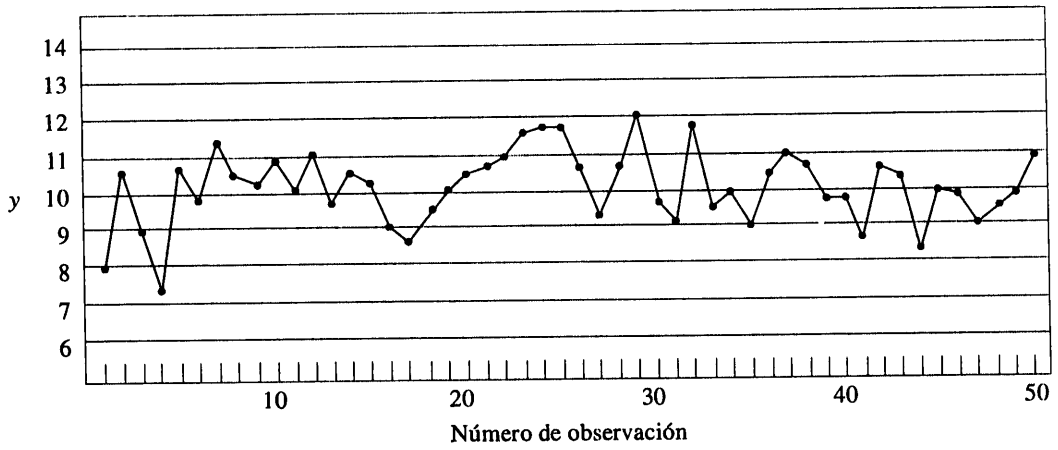


Figura 1-17*a* Datos del proceso que incluye únicamente perturbaciones aleatorias.

En la figura 1-18*a* se muestran los mismos datos de la figura 1-17*a*, excepto porque las mediciones después de la observación número 25 se incrementan dos unidades para simular el efecto de un corrimiento en la media del proceso. Cuando hay un corrimiento auténtico en la media de un proceso, un ajuste puede ser conveniente. En la figura 1-18*b* se muestran los datos obtenidos cuando se aplica un ajuste (un decremento de dos unidades) a la media después de que se detecta el corrimiento (en la observación 28). Nótese que con este ajuste disminuyen las desviaciones del valor que quiere alcanzarse.

La cuestión de cuándo aplicar un ajuste (y de qué magnitud) empieza con la comprensión de los tipos de variación que afectan un proceso. Un **diagrama de control** es un recurso invaluable para examinar la variabilidad en datos orientados en el tiempo. En la figura 1-19 se presenta el diagrama de control de los datos de la concentración del proceso químico de la figura 1-15. La *línea central* del diagrama de control es tan sólo el promedio de las mediciones de la concentración de las 20 primeras muestras ($\bar{x} = 91.5$ g/l) cuando el proceso es estable. El *límite superior de control* y el *límite inferior de control* son un par de límites derivados estadísticamente que reflejan la variabilidad inherente o natural del proceso. Estos límites se localizan tres desviaciones

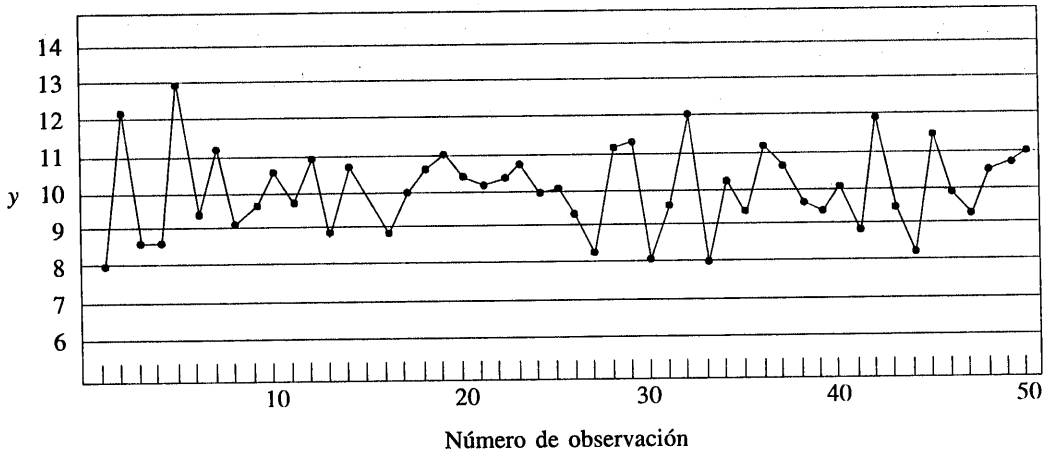


Figura 1-17*b* Los ajustes aplicados a las perturbaciones aleatorias introducen un sobrecontrol del proceso e incrementan las desviaciones respecto del valor que quiere alcanzarse.

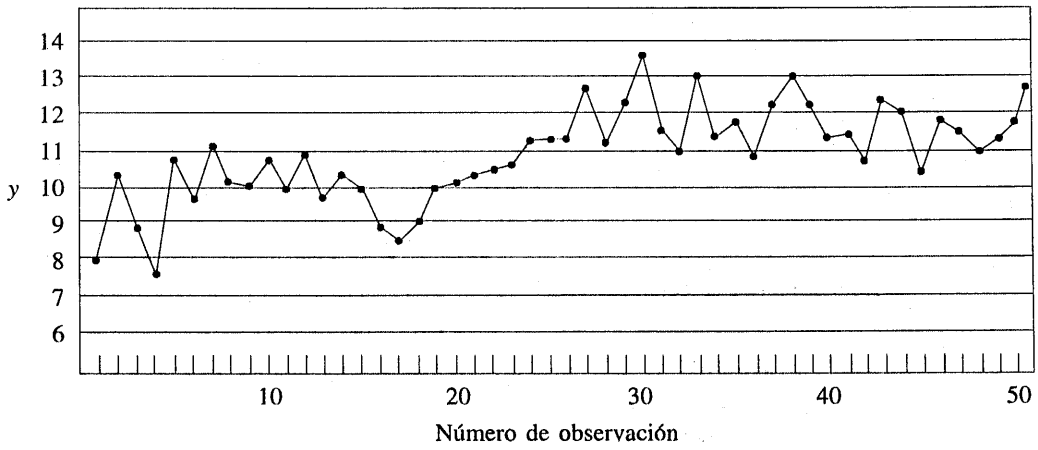


Figura 1-18*a* La media del proceso tiene un corrimiento (dos unidades hacia arriba) después de la observación número 25.

estándar de los valores de la concentración arriba y abajo de la línea central. Si el proceso está operando como debería, sin ninguna fuente externa de variabilidad presente en el sistema, las mediciones de la concentración deberán fluctuar alrededor de la línea central y la mayoría de ellas deberán quedar entre los límites de control.

En el diagrama de control de la figura 1-19, el marco de referencia visual que proporcionan la línea central y los límites de control indica que algún trastorno o perturbación ha incidido en el proceso cerca de la muestra 20, ya que todas las observaciones posteriores se encuentran abajo de la línea central y dos de ellas de hecho están abajo del límite inferior de control. Esta es una señal muy clara de que se requiere una acción correctiva en el proceso. Si es posible encontrar y eliminar la causa fundamental de este trastorno, el desempeño del proceso podrá mejorarse considerablemente.

Los diagramas de control son una aplicación muy importante de la estadística para monitorear, controlar y mejorar un proceso. La rama de la estadística que hace uso de los diagramas de control se llama **control estadístico de procesos** o **SPC** (Statistical Process Control). En el capítulo 15 se expondrán el SPC y los diagramas de control.

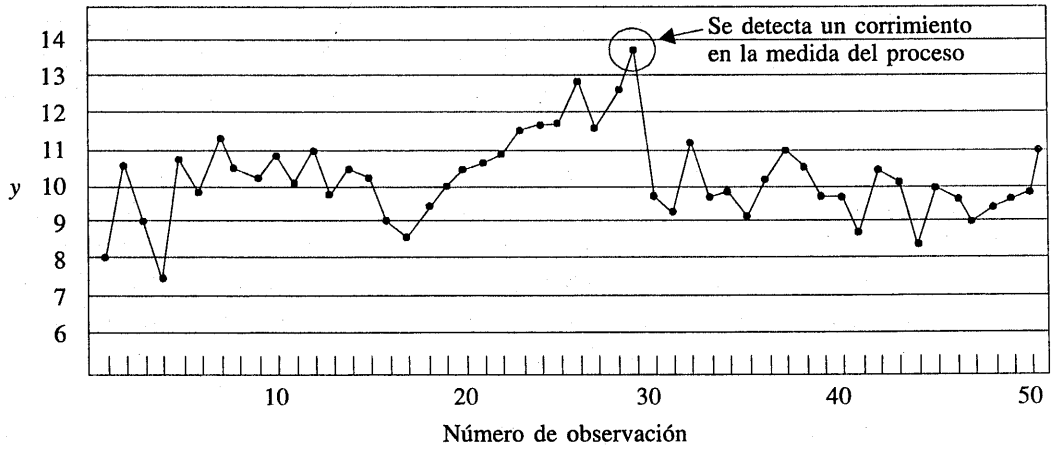


Figura 1-18*b* Se detecta un corrimiento en la media del proceso en la observación número 28 y las mediciones subsecuentes se decrementan dos unidades.

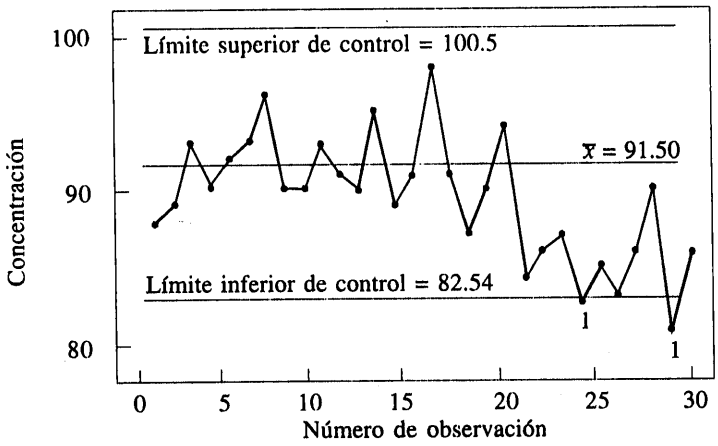


Figura 1-19 Diagrama de control para los datos de la concentración del proceso químico.

Ejercicios complementarios

1-9. Prevenir la propagación de las fisuras por fatiga en las estructuras aeronáuticas es un elemento importante de la seguridad en la aviación. Un estudio de ingeniería para investigar las fisuras por fatiga en $n = 9$ secciones de un ala sometidas a cargas por unidad de superficie cíclicas, reportó fisuras con las siguientes longitudes (en mm): 2.13, 2.96, 3.02, 1.82, 1.15, 1.37, 2.04, 2.47, 2.60.

a) Calcule la media muestral.

b) Calcule la varianza muestral y la desviación estándar muestral.

c) Haga un diagrama de puntos de los datos.

1-10. Considérense los datos de la intensidad solar del ejercicio 1-6. Haga un diagrama de puntos de estos datos. Indique dónde se localiza la media muestral en este diagrama. Proponga una interpretación práctica de la media muestral.

1-11. En el ejercicio 1-5 se describen los datos de un artículo de *Human Factors* sobre el acomodamiento visual obtenidos en un experimento con una pantalla catódica de alta resolución, CRT.

a) Construya un diagrama de puntos de estos datos.

b) En el artículo también se reportaron los datos de un segundo experimento utilizando una pantalla de baja resolución. Estos son 8.85, 35.80, 26.53, 64.63, 9.00, 15.38, 8.14 y 8.24. Haga un diagrama de puntos para esta segunda muestra y compárelo con el de la primera muestra. ¿Qué puede concluirse acerca de la resolución de la pantalla catódica en esta situación?

1-12. Un operador hace ocho mediciones del pH de una solución utilizando el mismo instrumento. Obtiene los siguientes datos: 7.15, 7.20, 7.18, 7.19, 7.21, 7.20, 7.16 y 7.18.

a) Calcule la media muestral.

b) Calcule la varianza muestral y la desviación estándar muestral.

c) ¿Cuáles son las principales fuentes de variabilidad en este experimento?

1-13. En un artículo de *Journal of Aircraft* se describe el cálculo de los coeficientes de retardo de la superficie aerodinámica NASA 0012. Se utilizaron algoritmos de cálculo diferentes para $M_\infty = 0.7$ con los resultados siguientes (los coeficientes de retardo de la superficie aerodinámica están en unidades de conteo de retardo aerodinámico; es decir, 1 conteo es equivalente a un coeficiente de retardo de la superficie aerodinámica de 0.0001): 79, 100, 74, 83, 81, 85, 82, 80 y 84. Calcule la media

muestral, la varianza muestral y la desviación estándar muestral, y haga un diagrama de puntos.

1-14. Los datos siguientes corresponden a las temperaturas conjuntas (°F) de los anillos sellos (*O-rings*) para cada encendido de prueba o lanzamiento real del motor del transbordador espacial (de *Presidential Commission on the Space Shuttle Challenger Accident*, vol. 1, pp. 129-131): 84, 49, 61, 40, 83, 67, 45, 66, 70, 69, 80, 58, 68, 60, 67, 72, 73, 70, 57, 63, 70, 78, 52, 67, 53, 67, 75, 61, 70, 81, 76, 79, 75, 76, 58, 31.

a) Calcule la media muestral y la desviación estándar muestral.

b) Construya un diagrama de puntos de los datos de las temperaturas.

d) Omita la observación más pequeña (31°F) y calcule de nuevo las cantidades del inciso a). Comente los resultados. ¿Qué tan "diferentes" son las otras temperaturas de este último valor?

EJERCICIOS PARA DESARROLLAR EL INTELECTO

1-15. Considere los datos de la superficie aerodinámica del ejercicio 1-13. Reste 30 de cada valor y después multiplique por 10 las cantidades resultantes. Calcule ahora s^2 para los nuevos datos. ¿Cómo se relaciona esta cantidad con la s^2 de los datos *originales?* Explique por qué.

1-16. Considere la cantidad $\sum_{i=1}^{n}(x_i - a)^2$. ¿Con qué valor de a se minimiza esta cantidad?

1-17. Usando los resultados del ejercicio 1-16, ¿cuál de las dos cantidades $\sum_{i=1}^{n}(x_i - \bar{x})^2$ y $\sum_{i=1}^{n}(x_i - \mu)^2$ será menor, considerando que $\bar{x} \neq \mu$?

1-18. **Codificación de los datos.** Sea $y_i = a + bx_i$, $i = 1, 2, \ldots, n$, donde a y b son constantes diferentes de cero. Encuentre la relación entre $\bar{x}$ y $\bar{y}$, y entre s_x y s_y.

1-19. Una muestra de mediciones de la temperatura en un horno produjo un promedio muestral (°F) de 835.00 y una desviación estándar muestral de 10.5. Utilizando los resultados del ejercicio 1-18, ¿cuáles son el promedio muestral y la desviación estándar muestral expresados en °C?

1-20. Considere la muestra $x_1, x_2, \ldots, x_n$ con media muestral $\bar{x}$ y desviación estándar muestral s. Sea $z_i = (x_i - \bar{x})/s$, $i = 1, 2, \ldots, n$. ¿Cuáles son los valores de la media muestral y la desviación estándar muestral de las z_i?

2 Resumen y presentación de datos

PLAN GENERAL DEL CAPÍTULO

2-1 LA IMPORTANCIA DEL RESUMEN Y LA PRESENTACIÓN DE DATOS

2-2 DIAGRAMAS DE TALLO Y HOJA

2-3 DISTRIBUCIONES DE FRECUENCIA E HISTOGRAMAS

2-4 GRÁFICAS DE CAJA

2-5 GRÁFICAS DE SERIES DE TIEMPO

2-1 LA IMPORTANCIA DEL RESUMEN Y LA PRESENTACIÓN DE DATOS

En el capítulo 1 se introdujeron la media muestral, la varianza muestral, la desviación estándar muestral y el diagrama de puntos como técnicas para resumir datos. Los resúmenes y las representaciones de datos bien elaborados son esenciales para el pensamiento estadístico correcto, ya que pueden enfocar al ingeniero en características importantes de los datos o proporcionar ideas acerca del modelo que debería emplearse para la solución del problema.

La computadora se ha convertido en una herramienta importante para la representación y el análisis de datos. Aun cuando muchas técnicas estadísticas sólo requieren una calculadora de mano, este enfoque puede requerir mucho tiempo y esfuerzo, mientras que una computadora llevará a cabo las tareas con mucho mayor eficiencia.

La mayoría de los análisis estadísticos se hacen utilizando una biblioteca de programas especializados para estadística. El usuario introduce los datos y después selecciona el tipo de análisis y el formato de la salida que desee. Se cuenta con paquetes de software de estadística tanto para mainframes como para computadoras personales. Entre los paquetes más populares y de uso más generalizado se encuentran el SAS (Statistical Analysis System), tanto para servido-

res como para computadoras personales (PC), y Statgraphics y Minitab para PC. A lo largo del libro se presentarán ejemplos de salidas de varios paquetes de estadística. No se discutirán los procedimientos particulares para introducir y editar los datos o para utilizar los comandos de los paquetes. El lector encontrará estos paquetes, u otros similares, en su escuela, así como la asesoría para aprender a usarlos.

2-2 DIAGRAMAS DE TALLO Y HOJA

El diagrama de puntos es una representación de datos útil para muestras pequeñas, hasta (digamos) unas 20 observaciones. Sin embargo, cuando el número de observaciones es más o menos grande, otras representaciones gráficas pueden resultar más convenientes.

Por ejemplo, considérense los datos de la tabla 2-1. Son datos de la resistencia a la compresión en libras por pulgada cuadrada (psi, *pounds per square inch*) de 80 ejemplares de una nueva aleación aluminio-litio sujeta a evaluación como posible material para elementos estructurales de aeronaves. Los datos se registraron en el orden de las pruebas y en este formato no comunican gran información sobre la resistencia a la compresión. No es sencillo responder a preguntas tales como: "¿Qué porcentaje de los ejemplares están abajo de 120 psi?". Debido al gran número de observaciones, construir el diagrama de puntos de estos datos resultaría un tanto infructuoso. Sin embargo, se cuenta con representaciones más eficaces para conjuntos grandes de datos.

Un **diagrama de tallo y hoja** es una forma adecuada de obtener una representación visual informativa de un conjunto de datos $x_1, x_2, ..., x_n$, donde cada número x_i tiene al menos dos dígitos. Para construir un diagrama de tallo y hoja, cada número x_i se divide en dos partes: un **tallo**, compuesto por uno o más de los primeros dígitos, y una **hoja**, compuesta por los dígitos restantes. Como ilustración, si los datos consisten en información del porcentaje entre 0 y 100 de piezas defectuosas en lotes de obleas de semiconductores, entonces el valor 76 puede dividirse en el tallo 7 y la hoja 6. En general, deberán elegirse relativamente pocos tallos en comparación con el número de observaciones. La mejor elección suele ser entre 5 y 20 tallos. Una vez que se ha elegido un conjunto de tallos, se enlistan en el margen izquierdo del diagrama. En seguida de cada tallo se enlistan todas las hojas correspondientes a los valores de los datos observados en el orden en que se van encontrando en el conjunto de datos.

Tabla 2-1 Resistencia a la compresión de 80 ejemplares de prueba de una aleación aluminio-litio

105	221	183	186	121	181	180	143
97	154	153	174	120	168	167	141
245	228	174	199	181	158	176	110
163	131	154	115	160	208	158	133
207	180	190	193	194	133	156	123
134	178	76	167	184	135	229	146
218	157	101	171	165	172	158	169
199	151	142	163	145	171	148	158
160	175	149	87	160	237	150	135
196	201	200	176	150	170	118	149

EJEMPLO 2-1

Para ilustrar la construcción de un diagrama de tallo y hoja, considérense los datos de la resistencia a la compresión de la tabla 2-1. Se seleccionarán como valores del tallo los números 7, 8, 9, ..., 24. El diagrama de tallo y hoja resultante se muestra en la figura 2-1. La última columna del diagrama es un conteo de la frecuencia del número de hojas asociadas con cada tallo. La inspección de esta representación revela de inmediato que la mayoría de las resistencias a la compresión se encuentran entre 110 y 200 psi y que un valor central se localiza en algún valor entre 150 y 160 psi. Además, las resistencias se distribuyen en forma aproximadamente simétrica alrededor del valor central. El diagrama de tallo y hoja permite determinar con rapidez algunas características importantes de los datos que no eran evidentes a primera vista en la representación original de la tabla 2-1.

En algunos conjuntos de datos puede resultar conveniente incluir más clases o tallos. Una forma de hacerlo sería modificando los tallos originales de la siguiente manera: el tallo 5 (por ejemplo) se divide en dos nuevos tallos, 5L y 5U. El tallo 5L tiene las hojas 0, 1, 2, 3 y 4, y el tallo 5U tiene las hojas 5, 6, 7, 8 y 9. De este modo se duplicará el número original de tallos. El número original de tallos podría incrementarse cuatro veces definiendo cinco nuevos tallos: 5z con las hojas 0 y 1, 5t con las hojas 2 y 3, 5f con las hojas 4 y 5, 5s con las hojas 6 y 7, y 5e con las hojas 8 y 9.

Tallo	Hoja	Frecuencia
7	6	1
8	7	1
9	7	1
10	5 1	2
11	5 8 0	3
12	1 0 3	3
13	4 1 3 5 3 5	6
14	2 9 5 8 3 1 6 9	8
15	4 7 1 3 4 0 8 8 6 8 0 8	12
16	3 0 7 3 0 5 0 8 7 9	10
17	8 5 4 4 1 6 2 1 0 6	10
18	0 3 6 1 4 1 0	7
19	9 6 0 9 3 4	6
20	7 1 0 8	4
21	8	1
22	1 8 9	3
23	7	1
24	5	1

Figura 2-1 Diagrama de tallo y hoja para los datos de la resistencia a la compresión de la tabla 2-1.

EJEMPLO 2-2

En la figura 2-2 se ilustra la representación de tallo y hoja de 25 observaciones de varios lotes de producción de un proceso químico. En la figura 2-2*a* se usan como tallos 6, 7, 8 y 9. Esto resulta en un número muy reducido de tallos y el diagrama de tallo y hoja no ofrece gran información acerca de los datos. En la figura 2-2*b* cada tallo se divide en dos partes, resultando una representación que muestra de manera más adecuada los datos. En la figura 2-2*c* se ilustra una representación de tallo y hoja con cada tallo dividido en cinco partes. Hay demasiados tallos en esta presentación, dando como resultado una representación que no dice gran cosa acerca de la forma de los datos.

En la figura 2-3 se muestra una representación de tallo y hoja creada con Minitab para los datos de la resistencia a la compresión de la tabla 2-1. El software utiliza los mismos tallos que en la figura 2-1. Nótese que la computadora ordena las hojas de menor a mayor en cada tallo. A esta forma de presentación suele llamarse **representación ordenada de tallo y hoja**. Por lo general esta ordenación se omite cuando la presentación se hace a mano debido a que puede tomar algún tiempo. La computadora agrega una columna a la izquierda de los tallos, en la cual se proporciona un conteo de las observaciones de cada tallo y los anteriores en la mitad superior de la representación, así como un conteo de las observaciones de cada tallo y los posteriores en la mitad inferior de la representación. En el tallo de en medio 16, la columna indica el número de observaciones incluidas en el mismo.

Tallo	Hoja
6	1 3 4 5 5 6
7	0 1 1 3 5 7 8 8 9
8	1 3 4 4 7 8 8
9	2 3 5

(*a*)

Tallo	Hoja
6L	1 3 4
6U	5 5 6
7L	0 1 1 3
7U	5 7 8 8 9
8L	1 3 4 4
8U	7 8 8
9L	2 3
9U	5

(*b*)

Tallo	Hoja
6z	1
6t	3
6f	4 5 5
6s	6
6e	
7z	0 1 1
7t	3
7f	5
7s	7
7e	8 8 9
8z	1
8t	3
8f	4 4
8s	7
8e	8 8
9z	
9t	2 3
9f	5
9s	
9e	

(*c*)

Figura 2-2 Representaciones de tallo y hoja para el ejemplo 2-2.

Representación del diagrama de tallo y hoja

Tallo y hoja para resistencia

N = 80 Unidad de hoja = 1.0

```
 1  7    6
 2  8    7
 3  9    7
 5 10    1 5
 8 11    0 5 8
11 12    0 1 3
17 13    1 3 3 4 5 5
25 14    1 2 3 5 6 8 9 9
37 15    0 0 1 3 4 4 6 7 8 8 8 8
(10) 16  0 0 0 3 3 5 7 7 8 9
33 17    0 1 1 2 4 4 5 6 6 8
23 18    0 0 1 1 3 4 6
16 19    0 3 4 6 9 9
10 20    0 1 7 8
 6 21    8
 5 22    1 8 9
 2 23    7
 1 24    5
```

Figura 2-3 Diagrama de tallo y hoja de Minitab.

La representación ordenada de tallo y hoja hace relativamente sencillo determinar características de los datos, tales como los percentiles, los cuartiles y la mediana. La **mediana** muestral es una medida de tendencia central que divide los datos en dos partes iguales, la mitad que queda abajo de la mediana y la mitad que queda arriba de ella. Si el número de observaciones es par, la mediana está a la mitad de los dos valores centrales. En la figura 2-3 se encuentra que los valores 40 y 41 de la resistencia son 160 y 163, por lo que la mediana es $(160 + 163)/2 = 161.5$. Si el número de observaciones es impar, la mediana es el valor central. La **moda** muestral es el valor de los datos que ocurre con mayor frecuencia. La figura 2-3 indica que la moda es 158; este valor ocurre cuatro veces y ningún otro valor se repite con tanta frecuencia en la muestra.

Los datos también pueden dividirse en más de dos partes. Cuando un conjunto ordenado de datos se divide en cuatro partes iguales, los puntos de división se denominan **cuartiles**. El *primer cuartil* o *cuartil inferior*, q_1, es un valor que tiene aproximadamente una cuarta parte (25%) de la observaciones abajo de él y aproximadamente 75% de las observaciones arriba. El *segundo cuartil*, q_2, tiene aproximadamente la mitad (50%) de las observaciones abajo de su valor. El segundo cuartil es exactamente igual a la mediana. El *tercer cuartil*, q_3, tiene aproximadamente tres cuartas partes (75%) de las observaciones abajo de su valor. Como en el caso de la mediana, los cuartiles pueden no ser únicos. Los datos de la resistencia a la compresión de la figura 2-3 contienen $n = 80$ observaciones. El software Minitab calcula los cuartiles primero y tercero como las observaciones ordenadas $(n + 1)/4$ y $3(n + 1)/4$ y hace las interpolaciones necesarias. Por ejemplo, $(80 + 1)/4 = 20.25$ y $3(80 + 1)/4 = 60.75$. Por lo tanto, Minitab hace la interpolación entre las observaciones ordenadas 20 y la 21 para obtener $q_1 = 143.50$ y entre las observaciones

Tabla 2-2 Estadísticas de resumen para los datos de la resistencia a la compresión usando Minitab

Variable	N	Media	Mediana	Desv. Est.	Media de EE
	80	162.66	161.50	33.77	3.78
	Mín.	Máx.	Q1	Q3	
	76.00	245.00	143.50	181.00	

60 y la 61 para obtener $q_3 = 181.00$. En general, el **percentil** $100k$-ésimo es un valor de los datos tal que aproximadamente el $100k\%$ de las observaciones está en este valor o abajo del mismo y aproximadamente el $100(1 - k)\%$ de ellas se encuentra arriba de él. Por último, puede usarse un **rango intercuartílico**, definido como $\text{IQR} = q_3 - q_1$, como medida de variabilidad. El rango intercuartílico es menos sensible a los valores extremos de la muestra que el rango muestral ordinario.

Muchos paquetes de software de estadística proporcionan resúmenes de datos que incluyen estas cantidades. En la tabla 2-2 se muestra la salida de Minitab que se obtiene para los datos de la resistencia a la compresión de la tabla 2-1.

EJERCICIOS DE LA SECCIÓN 2-2

2-1. Un artículo de *Technometrics* (vol. 19, p. 425) presenta los siguientes datos sobre el octanaje del combustible para motores de varias marcas de gasolina:

88.5	87.7	83.4	86.7	87.5
94.7	91.1	91.0	94.2	87.8
84.3	86.7	88.2	90.8	88.3
90.1	93.4	88.5	90.1	89.2
89.0	96.1	93.3	91.8	92.3
89.8	89.6	87.4	88.4	88.9
91.6	90.4	91.1	92.6	89.8
90.3	91.6	90.5	93.7	92.7
90.0	90.7	100.3	96.5	93.3
91.5	88.6	87.6	84.3	86.7
89.9	88.3	92.7	93.2	91.0
98.8	94.2	87.9	88.6	90.9
88.3	85.3	93.0	88.7	89.9
90.4	90.1	94.4	92.7	91.8
91.2	89.3	90.4	89.3	89.7
90.6	91.1	91.2	91.0	92.2
92.2	92.2			

Construya una representación de tallo y hoja para estos datos.

2-2. Los siguientes datos corresponden al número de ciclos para producir la falla de una muestra de ensayo de aluminio sometida a un esfuerzo alternado repetido de 21 000 psi, 18 ciclos por segundo:

1 115	1 567	1 223	1 782	1 055
1 310	1 883	375	1 522	1 764
1 540	1 203	2 265	1 792	1 330
1 502	1 270	1 910	1 000	1 608
1 258	1 015	1 018	1 820	1 535
1 315	845	1 452	1 940	1 781
1 085	1 674	1 890	1 120	1 750
798	1 016	2 100	910	1 501
1 020	1 102	1 594	1 730	1 238
865	1 605	2 023	1 102	990
2 130	706	1 315	1 578	1 468
1 421	2 215	1 269	758	1 512
1 109	785	1 260	1 416	1 750
1 481	885	1 888	1 560	1 642

a) Construya una representación de tallo y hoja de estos datos.

b) ¿Parece factible que una muestra de ensayo "sobreviva" más allá de 2 000 ciclos? Justifique su respuesta.

2-3. Se presenta a continuación el porcentaje de algodón en el material usado para fabricar camisas para caballero. Construya una representación de tallo y hoja de los datos.

34.2	33.6	33.8	34.7
33.1	34.7	34.2	33.6
34.5	35.0	33.4	32.5
35.6	35.4	34.7	34.1
36.3	36.2	34.6	35.1
35.1	36.8	35.2	36.8
34.7	35.1	35.0	37.9
33.6	35.3	34.9	36.4
37.8	32.6	35.8	34.6
36.6	33.1	37.6	33.6
35.4	34.6	37.3	34.1
34.6	35.9	34.6	34.7
33.8	34.7	35.5	35.7
37.1	33.6	32.8	36.8
34.0	32.9	32.1	34.3
34.1	33.5	34.5	32.7

2-4. Los datos que se muestran a continuación representan el rendimiento de 90 lotes consecutivos de sustrato cerámico al que se le ha aplicado un recubrimiento metálico mediante un proceso de deposición de vapor. Construya una representación de tallo y hoja para estos datos.

94.1	87.3	94.1	92.4	84.6	85.4
93.2	84.1	92.1	90.6	83.6	86.6
90.6	90.1	96.4	89.1	85.4	91.7
91.4	95.2	88.2	88.8	89.7	87.5
88.2	86.1	86.4	86.4	87.6	84.2
86.1	94.3	85.0	85.1	85.1	85.1
95.1	93.2	84.9	84.0	89.6	90.5
90.0	86.7	78.3	93.7	90.0	95.6
92.4	83.0	89.6	87.7	90.1	88.3
87.3	95.3	90.3	90.6	94.3	84.1
86.6	94.1	93.1	89.4	97.3	83.7
91.2	97.8	94.6	88.6	96.8	82.9
86.1	93.1	96.3	84.1	94.4	87.3
90.4	86.4	94.7	82.6	96.1	86.4
89.1	87.6	91.1	83.1	98.0	84.5

2-5. Encuentre la mediana y los cuartiles para los datos del octanaje del combustible para motor del ejercicio 2-1.

2-6. Encuentre la mediana y los cuartiles para los datos de falla de las muestras de ensayo del ejercicio 2-2.

2-7. Encuentre la mediana, la moda y el promedio muestral de los datos del ejercicio 2-3. Explique la forma en que estas tres medidas de localización describen diferentes características de los datos.

2-8. Encuentre la mediana y los cuartiles para los datos de rendimiento del ejercicio 2-4.

2-3 DISTRIBUCIONES DE FRECUENCIA E HISTOGRAMAS

Una **distribución de frecuencia** es un resumen de datos más compacto que un diagrama de tallo y hoja. Para construir una distribución de frecuencia, el rango de los datos debe dividirse en intervalos, a los que acostumbra llamarse **intervalos de clase** o **celdas**. De ser posible, los intervalos de clase deberán ser de la misma anchura a fin de reforzar la información visual de la distribución de frecuencia. La selección del número de intervalos de clase deberá hacerse con criterio para que pueda desarrollarse una representación razonable. El número de intervalos de clase depende del número de observaciones y del grado de dispersión de los datos. Una distribución de frecuencia que emplea muy pocos intervalos de clase o demasiados no será informativa. En general, se encuentra que entre 5 y 20 intervalos de clase son satisfactorios en la mayoría de los casos y que el número de intervalos de clase deberá incrementarse con n. Elegir un **número de intervalos de clase** aproximadamente igual a la raíz cuadrada del número de observaciones suele dar buenos resultados en la práctica.

En la tabla 2-3 se muestra la distribución de frecuencia para los datos de la resistencia a la compresión de la tabla 2-1. Puesto que el conjunto de datos contiene 80 observaciones, y como $\sqrt{80} \cong 9$, se supone que entre ocho y nueve celdas proporcionarán una distribución de frecuen-

Tabla 2-3 Distribución de frecuencia para los datos de la resistencia a la compresión de la tabla 2-1

Intervalo de clase (psi)	Conteo	Frecuencia	Frecuencia relativa	Frecuencia relativa acumulada
$70 \leq x < 90$	II	2	0.0250	0.0250
$90 \leq x < 110$	III	3	0.0375	0.0625
$110 \leq x < 130$	~~IIII~~ I	6	0.0750	0.1375
$130 \leq x < 150$	~~IIII~~ ~~IIII~~ IIII	14	0.1750	0.3125
$150 \leq x < 170$	~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~ II	22	0.2750	0.5875
$170 \leq x < 190$	~~IIII~~ ~~IIII~~ ~~IIII~~ II	17	0.2125	0.8000
$190 \leq x < 210$	~~IIII~~ ~~IIII~~	10	0.1250	0.9250
$210 \leq x < 230$	IIII	4	0.0500	0.9750
$230 \leq x < 250$	II	2	0.0250	1.0000

cia satisfactoria. Los valores más grande y más pequeño son 245 y 76, respectivamente, por lo que los intervalos de clase deben cubrir un rango de al menos 245 – 76 = 169 unidades en la escala psi. Si se desea que el límite inferior del primer intervalo empiece un poco antes del valor más pequeño y que el límite superior del último intervalo de clase esté un poco después del valor más grande, entonces la distribución de frecuencia podría empezar en 70 y terminar en 250. Éste es un intervalo o rango de 180 unidades psi. Nueve intervalos de clase, cada uno de 20 psi de anchura, producen una distribución de frecuencia razonable, por lo que la distribución de frecuencia de la tabla 2-3 se basa en nueve intervalos de clase.

La cuarta columna de la tabla 2-3 contiene una **distribución de frecuencia relativa**. Las frecuencias relativas se determinan dividiendo la frecuencia observada en cada intervalo de clase por el número total de observaciones. En la última columna de la tabla 2-3 se expresan las frecuencias relativas sobre una base acumulada. La interpretación de una distribución de frecuencia suele ser más sencilla que la de una tabla de datos. Por ejemplo, en la tabla 2-3 es muy fácil ver que la mayoría de los ejemplares para ensayo tienen resistencias a la compresión entre 130 y 190 psi y que el 97.5% de los ejemplares para ensayo están abajo de 230 psi.

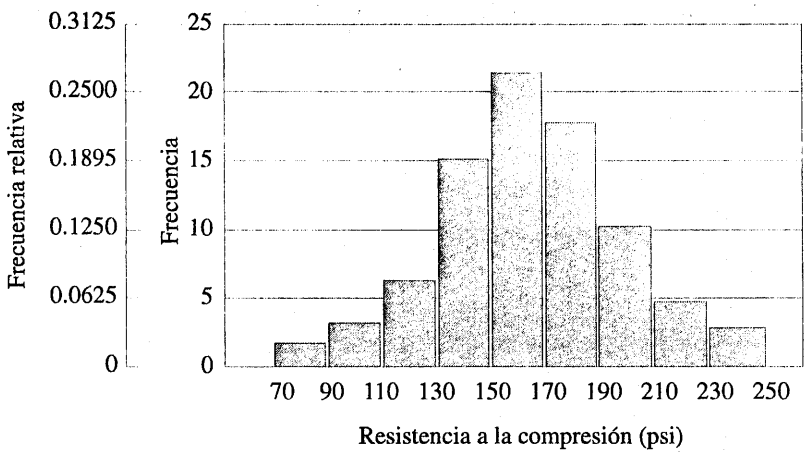


Figura 2-4 Histograma de la resistencia a la compresión de 80 ejemplares de prueba de la aleación litio-aluminio.

También resulta conveniente presentar la distribución de frecuencia en forma gráfica, como se muestra en la figura 2-4. A esta representación se le llama **histograma**. Para hacer un histograma, se usa el eje horizontal para representar la escala de medición y para trazar los límites de los intervalos de clase. El eje vertical representa la escala de la frecuencia (o de la frecuencia relativa). Si los intervalos de clase tienen el mismo ancho, entonces las alturas de los rectángulos trazados en los histogramas son proporcionales a las frecuencias. Si los intervalos de clase son de anchuras diferentes, entonces se acostumbra trazar rectángulos cuyas áreas sean proporcionales a las frecuencias. Sin embargo, la interpretación de los histogramas se facilita cuando los intervalos de clase tienen la misma anchura. El histograma, al igual que el diagrama de tallo y hoja, proporciona una impresión visual de la forma que asume la distribución de las mediciones, así como información acerca de la separación o dispersión de los datos. Nótese la distribución simétrica, en forma de campana, de las mediciones de la resistencia a la compresión de la figura 2-4.

Al pasar de los datos originales o del diagrama de tallo y hoja a una distribución de frecuencia o a un histograma, se pierde cierta información porque ya no se tienen las observaciones individuales. Esta pérdida de información suele ser reducida en comparación con la brevedad y la facilidad de interpretación que se gana al utilizar la distribución de frecuencia y el histograma.

En la figura 2-5 se ilustra el histograma de Minitab para los datos de la resistencia a la compresión. En este histograma se utilizaron los ajustes "por omisión" del programa, los cuales dieron como resultado 17 intervalos de clase. Se ha señalado ya que los histogramas pueden ser relativamente sensibles al número de intervalos de clase y a su anchura. Para conjuntos pequeños de datos, la apariencia de los histogramas puede sufrir cambios radicales si se modifica el número y/o la anchura de los intervalos de clase. Los histogramas son más estables para conjuntos grandes de datos, de preferencia con un tamaño de 75 a 100 o más. En la figura 2-6 se muestra el histograma de Minitab para los datos de la resistencia a la compresión con nueve intervalos de clase, el cual es similar al histograma original mostrado en la figura 2-4. Puesto que el número de observaciones es moderadamente grande ($n = 80$), la elección del número de intervalos de clase no es de particular importancia y las figuras 2-5 y 2-6 comunican información similar.

En la figura 2-7 se muestra una variante del histograma disponible en Minitab, la **gráfica de frecuencias acumuladas**. En esta gráfica, la altura de cada barra es el número total de observaciones que son menores o iguales que el límite superior del intervalo de clase. Las distribuciones

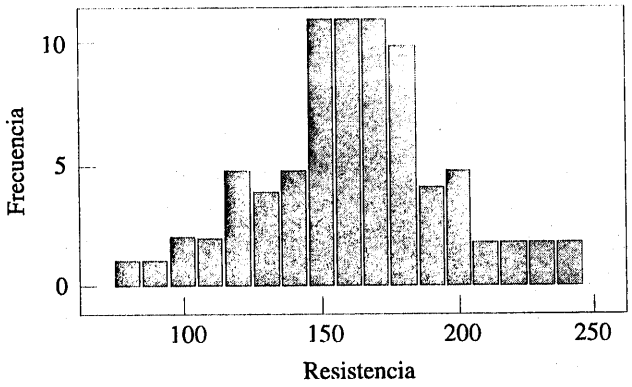


Figura 2-5 Histograma de Minitab con 17 intervalos para los datos de la resistencia a la compresión.

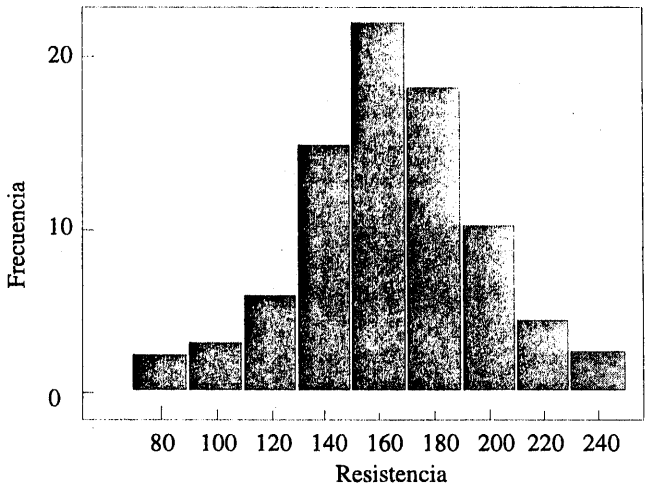


Figura 2-6 Histograma de Minitab con nueve intervalos para los datos de la resistencia a la comprensión.

acumuladas también son útiles en la interpretación de datos; por ejemplo, en la figura 2-7 puede leerse directamente que hay cerca de 70 observaciones menores o iguales que 200 psi.

Los histogramas son más efectivos como representaciones de datos para muestras relativamente grandes, por ejemplo $n \geq 75$ a 100 o más. Cuando el tamaño de la muestra es grande, el histograma puede proporcionar un indicador confiable de la **forma** general de la población de mediciones de la que se tomó la muestra. En la figura 2-8 se presentan tres casos. En general, si los datos son simétricos, como en la figura 2-8*b*, entonces la media y la mediana coinciden. Si, además, los datos tienen una sola moda (se dice entonces que los datos son *unimodales*), entonces la media, la mediana y la moda coinciden. Si los datos están *sesgados* (son asimétricos, con una cola larga a uno de los lados), como en las figuras 2-8*a* y 2-8*c*, entonces la media, la mediana y la moda no coinciden. Por lo general se encuentra que moda < mediana < media, si la distribución está sesgada a la derecha, en tanto que moda > mediana > media, si la distribución está sesgada a la izquierda.

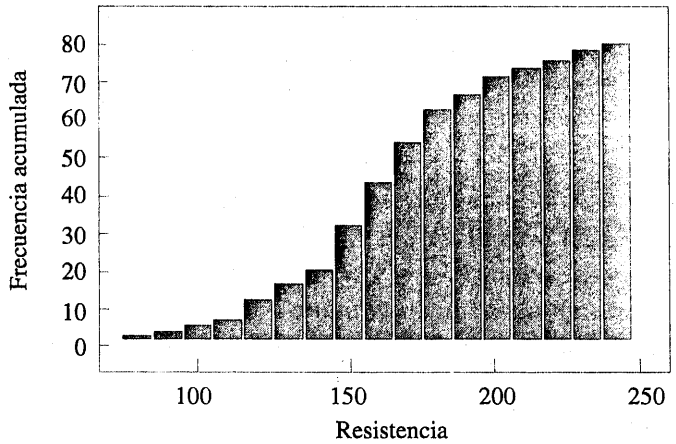


Figura 2-7 Gráfica de Minitab de la distribución acumulada de los datos de la resistencia a la comprensión.

Las distribuciones de frecuencias y los histogramas también pueden usarse con datos cualitativos o datos categóricos. En algunas aplicaciones habrá una ordenación natural de las categorías (como estudiante de primaria, secundaria, bachillerato y profesional), en tanto que en otras el orden de las categorías será arbitrario (como masculino y femenino). Cuando se usan datos categóricos, los intervalos de clase deberán trazarse con la misma anchura.

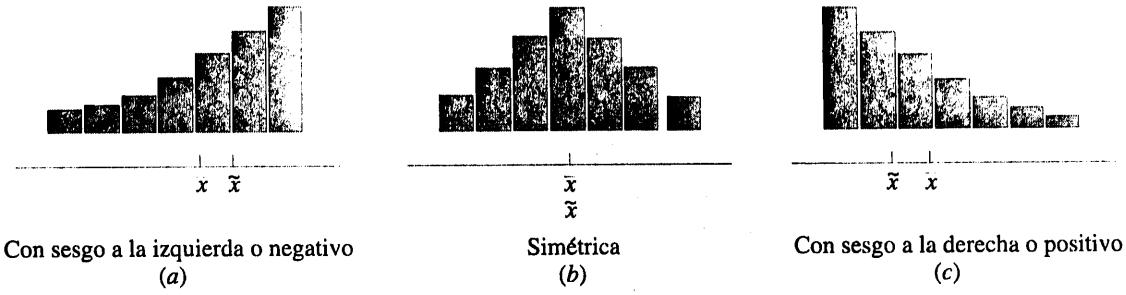


Figura 2-8 Histogramas de distribuciones simétricas y sesgadas.

EJEMPLO 2-3

En la figura 2-9 se presenta la producción de aviones de pasajeros de Boeing Company en 1985. Obsérvese que el 737 fue el modelo más popular, seguido por los modelos 757, 747, 767 y 707.

Un histograma de ocurrencia por categoría (en el que las categorías están ordenadas por el número de ocurrencias) en ocasiones se denomina **diagrama de Pareto** (véase el ejercicio 2-14).

Esta sección se ha concentrado en los métodos descriptivos para la situación en que cada observación de un conjunto de datos es un solo número o pertenece a una sola categoría. Sin embargo, muchas veces se trabaja con datos en los que cada observación consta de varias mediciones. Por ejemplo, en un estudio del rendimiento de combustible, cada observación podría consistir en una medición de los kilómetros recorridos por litro, del cilindraje del motor del vehículo, de los caballos de fuerza del motor, del peso del vehículo y de la longitud del mismo. Este es un ejemplo de **datos con multivariables**. En capítulos posteriores se discutirá el análisis de este tipo de datos.

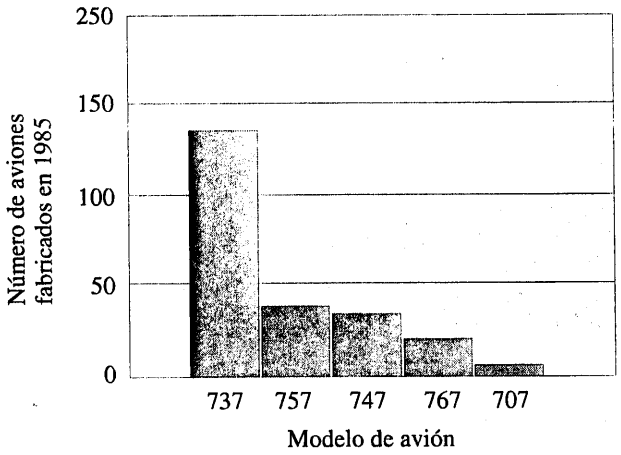


Figura 2-9 Producción de aviones en 1985. (*Fuente*: Boeing Company.)

EJERCICIOS DE LA SECCIÓN 2-3

2-9. Construya una distribución de frecuencia y un histograma para los datos del octanaje del combustible para motor del ejercicio 2-1. Use ocho intervalos de clase.

2-10. Construya una distribución de frecuencia y un histograma utilizando los datos de falla de las muestras de ensayo del ejercicio 2-2.

2-11. Construya una distribución de frecuencia y un histograma para los datos del contenido de algodón del ejercicio 2-3.

2-12. Construya una distribución de frecuencia y un histograma para los datos de rendimiento del ejercicio 2-4.

2-13. Construya una distribución de frecuencia y un histograma con 16 intervalos de clase para los datos del octanaje del combustible para motor del ejercicio 2-1. Compre su forma con la del histograma con ocho intervalos de clase del ejercicio 2-9. ¿Presentan información similar ambos histogramas?

2-14. **El diagrama de Pareto**. Una importante variación de un histograma para datos categóricos es el diagrama de Pareto. Este diagrama es de uso generalizado en los esfuerzos para mejorar la calidad, donde las categorías suelen representar diferentes tipos de defectos, condiciones de falla o problemas en productos/procesos. Las categorías se ordenan de tal modo que la categoría con la frecuencia más alta esté a la izquierda, seguida de la categoría con la segunda frecuencia más alta, y así sucesivamente. Estos diagramas deben su nombre al economista italiano V. Pareto y generalmente exhiben la "ley de Pareto"; es decir, la mayoría de los defectos pueden explicarse por unas cuantas categorías. Suponga que se obtuvo la siguiente información sobre los defectos estructurales en las puertas de un automóvil: abolladuras, 4, perforaciones, 4; partes ensambladas fuera de secuencia, 6; partes subrecortadas, 21; agujeros faltantes/ranuras, 8; partes no lubricadas, 5; partes mal perfiladas, 30; y partes sin esmerilar, 3. Construya e interprete un diagrama de Pareto.

2-4 GRÁFICAS DE CAJA

La representación de tallo y hoja y el histograma proporcionan impresiones visuales generales acerca de un conjunto de datos, en tanto que las cantidades numéricas tales como $\bar{x}$ o s proporcionan información acerca de una característica particular de los datos. La **gráfica de caja** es una representación gráfica que describe simultáneamente varias características importantes de un conjunto de datos, como el centro, la dispersión, la desviación de la simetría y la identificación de observaciones que caen inusualmente lejos del grueso de los datos. (A estas observaciones se les llama "puntos atípicos".)

En este tipo de gráfica se representan los tres cuartiles y los datos mínimo y máximo en una caja rectangular alineada en sentido horizontal o vertical. La caja abarca el rango intercuartílico, con el borde izquierdo (o el inferior) en el primer cuartil, q_1, y el borde derecho (o el superior) en el tercer cuartil, q_3. Se traza una línea a través de la caja en el segundo cuartil (que es el percentil 50 o la mediana), $q_2 = \tilde{x}$. Una línea, o **bigote**, se extiende desde cada extremo de la caja. El bigote bajo o izquierdo es una línea que va del primer cuartil al punto correspondiente al menor de los datos dentro de 1.5 rangos intercuartílicos a partir del primer cuartil. El bigote superior o derecho es una línea que va del tercer cuartil al punto correspondiente al mayor de los datos dentro de 1.5 rangos intercuartílicos a partir del tercer cuartil. Los datos que se encuentran alejados de la caja más allá de los bigotes se grafican como puntos individuales. A un punto situado después de un bigote, pero a menos de 3 rangos intercuartílicos del borde de la caja, se le llama **punto atípico**. A un punto situado a más de 3 rangos intercuartílicos del borde de la caja

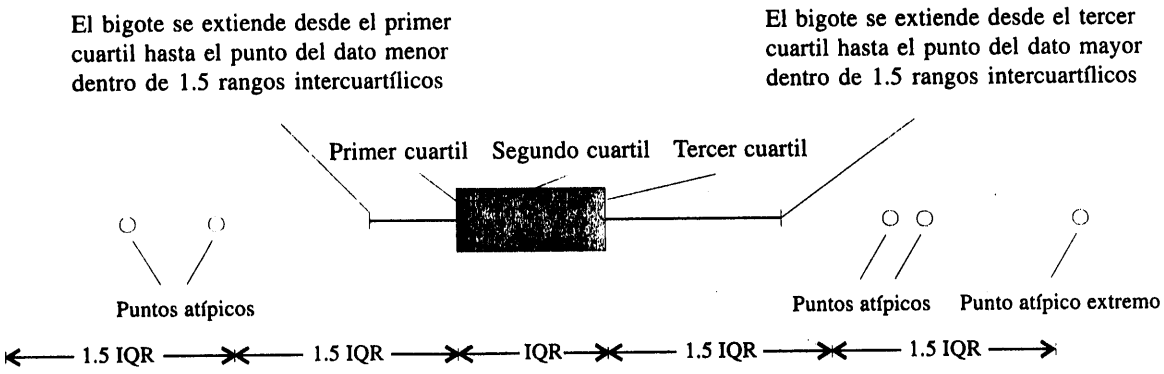


Figura 2-10 Descripción de una gráfica de caja.

se le llama **punto atípico extremo** (vease la figura 2-10). Ocasionalmente se usan símbolos diferentes, tales como círculos vacíos o rellenos, para identificar las dos clases de puntos atípicos. En ocasiones las gráficas de cajas reciben el nombre de *gráficas de cajas y bigotes.*

En la figura 2-11 se muestra la gráfica de caja de Minitab para los datos de la resistencia a la compresión de la aleación presentados en la tabla 2-1. Esta gráfica de caja indica que la distribución de las resistencias a la compresión es bastante simétrica alrededor del valor central, ya que tanto los bigotes izquierdo y derecho como las longitudes de las cajas izquierda y derecha alrededor de la mediana son aproximadamente iguales. Hay también dos puntos ligeramente atípicos en cada extremo de los datos.

Las gráficas de cajas son muy útiles en la comparación gráfica entre conjuntos de datos, ya que tienen un impacto visual inmediato y son fáciles de entender. Por ejemplo, en la figura 2-12 se muestran las gráficas de cajas comparativas para un índice de calidad en la fabricación de dispositivos semiconductores en tres plantas de manufactura. La inspección de esta representación revela que hay demasiada variabilidad en la planta 2 y que las plantas 2 y 3 necesitan elevar el desempeño de su índice de calidad.

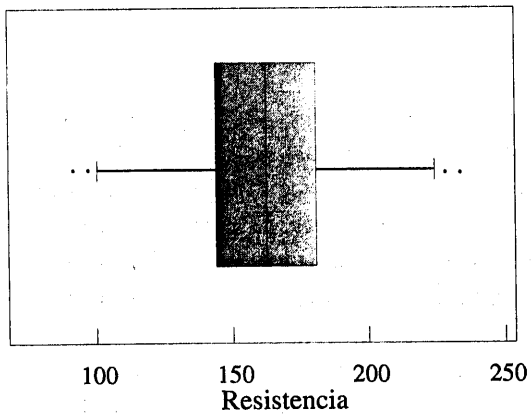


Figura 2-11 Gráfica de caja para los datos de la resistencia a la compresión de la tabla 2-1.

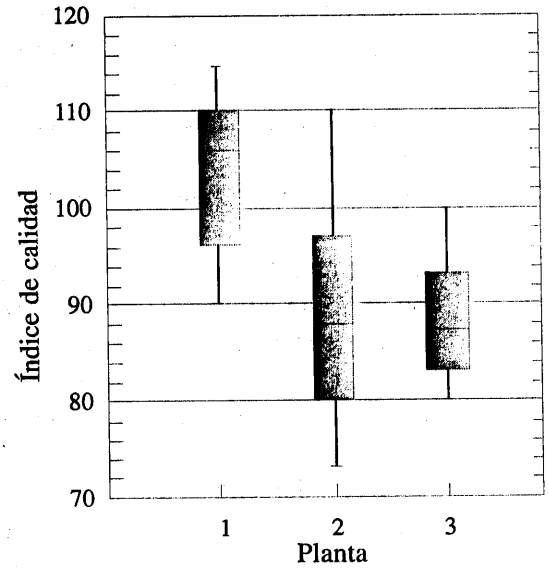


Figura 2-12 Gráfica de caja comparativas de un índice de calidad en tres plantas.

EJERCICIOS DE LA SECCIÓN 2-4

2-15. En el ejercicio 1-14 se presentaron las temperaturas conjuntas de los anillos sellos (*O-rings*) (°F) para cada encendido de prueba o lanzamiento real del motor del transbordador espacial. En ese ejercicio se pidió encontrar la media muestral y la desviación estándar muestral de la temperatura.

a) Encuentre los cuartiles superior e inferior de la temperatura.
b) Encuentre la mediana.
c) Omita la observación menor (31° F) y calcule de nuevo las cantidades de los incisos a) y b). Comente los resultados. ¿Qué tan "diferentes" son las demás temperaturas de este valor más pequeño?
d) Construya un diagrama de caja de los datos y comente la posible presencia de puntos apartados.

2-16. En un artículo de *Transactions of the Institution of Chemical Engineers* (vol. 34 , pp. 280-293) se presentaron los datos de un experimento que investigaba el efecto de varias variables del proceso sobre la oxidación en la fase de vapor de la naftalina. Una muestra de la conversión porcentual en moles de la naftalina en anhídrido maleico es: 4.2, 4.7, 4.7, 5.0, 3.8, 3.6, 3.0, 5.1, 3.1, 3.8, 4.8, 4.0, 5.2, 4.3, 2.8, 2.0, 2.8, 3.3, 4.8, 5.0.

a) Calcule la media muestral.
b) Calcule la varianza muestral y la desviación estándar muestral.
c) Construya una gráfica de caja de los datos.

2-17. Un fabricante de gasolina está investigando el "tiempo de ignición en el arranque en frío" de un motor de automóvil. Se obtuvieron los siguientes tiempos (en segundos) para un vehículo de prueba: 1.75, 1.92, 2.62, 2.35, 3.09, 3.15, 2.53, 1.91.

a) Calcule la media muestral y la desviación estándar muestral.
b) Construya una gráfica de caja de los datos.

2-18. Las nueve mediciones siguientes corresponden a las temperaturas de un horno registradas en lotes sucesivos de un proceso de fabricación de semiconductores (las unidades son °F): 953, 950, 948, 955, 951, 949, 957, 954, 955.

a) Calcule la media muestral, la varianza muestral y la desviación estándar muestral.
b) Encuentre la mediana. ¿Cuánto podría incrementarse la temperatura más alta sin que se altere el valor de la mediana?
c) Construya una gráfica de caja de los datos.

2-19. En el ejercicio 1-13 se presentan los coeficientes de retardo de la superficie aerodinámica NASA 0012. Se pidió ahí calcular la media muestral, la varianza muestral y la desviación estándar muestral de esos coeficientes.

a) Encuentre los cuartiles superior e inferior de los coeficientes de retardo.
b) Construya una gráfica de caja de los datos.
c) Omita la observación más grande (100) y repita los incisos a) y b) comente los resultados.

2-20. Los datos siguientes corresponden a las temperaturas del chorro de descarga de unas instalaciones de tratamiento de aguas negras en días consecutivos:

43	47	51	48	52	50	46	49
45	52	46	51	44	49	46	51
49	45	44	50	48	50	49	50

a) Calcule la media muestral y la mediana.
b) Calcule la varianza muestral y la desviación estándar muestral.
c) Construya una gráfica de caja de los datos y comente la información de esta representación.
d) Encuentre los percentiles 5 y 95 de la temperatura.

2-5 GRÁFICAS DE SERIES DE TIEMPO

Las representaciones gráficas que se han considerado hasta este punto —como los histogramas, los diagramas de tallo y hoja y las gráficas de caja— son métodos visuales muy útiles para mostrar la variabilidad de los datos. Sin embargo, en la sección 1-5 se señaló que el tiempo es un factor importante que contribuye a la variabilidad de los datos y en estos métodos gráficos no se toma en consideración este hecho. Una **serie de tiempo** o **secuencia cronológica** es un conjunto de datos en el que las observaciones se registran en el orden en que ocurren. Una **gráfica de series de tiempo** es una gráfica en la que el eje vertical denota el valor observado de la variable (por ejemplo, x) y el eje horizontal denota el tiempo (que podría estar en minutos, días, años, etcétera). Cuando las mediciones se grafican como una serie de tiempo, con frecuencia se observan tendencias, ciclos u otras características generales de los datos que de otro modo no podrían percibirse.

Por ejemplo, considérese la figura 2-13*a*, donde se presenta una gráfica de series de tiempo de las ventas anuales de una compañía durante los últimos 10 años. La impresión general de esta representación es que las ventas muestran una **tendencia** ascendente. Hay cierta variabilidad en esta tendencia, con las ventas de algunos años superando las del año anterior en tanto que las ventas de otros años son menores. En la figura 2-13*b* se muestran las ventas de los tres últimos años reportadas por trimestre. Esta gráfica muestra con claridad que las ventas anuales de este

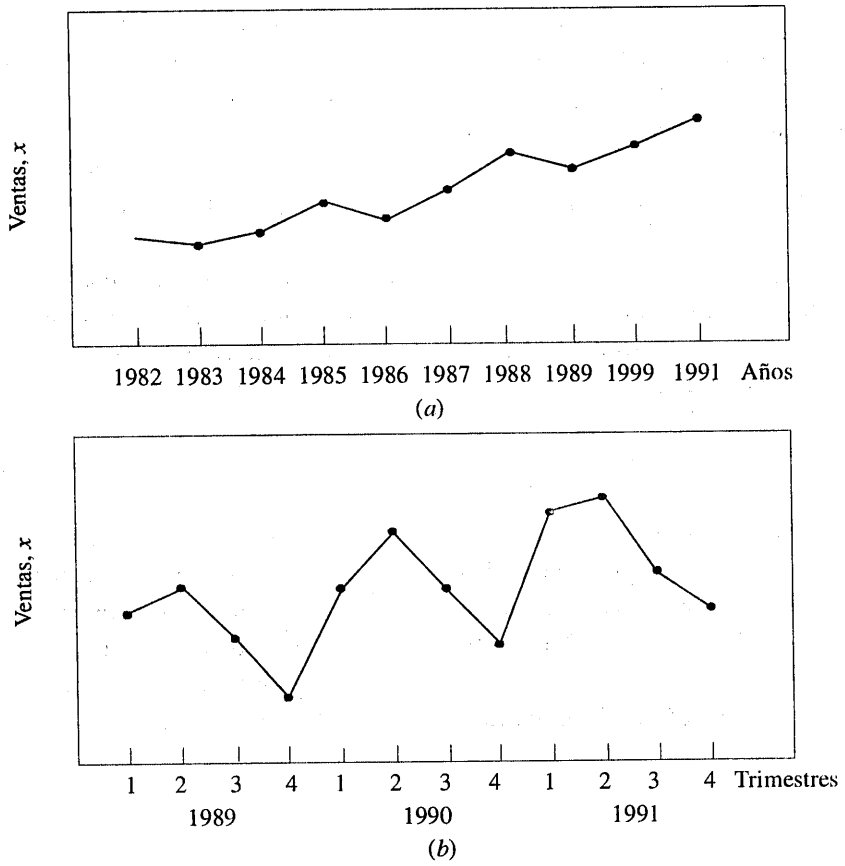


Figura 2-13 Ventas de una compañía por año (*a*) y por trimestre (*b*).

negocio presentan una variabilidad **cíclica** por trimestre, siendo en general las ventas en el primer y el segundo trimestre más altas que las ventas en el tercer y el cuarto trimestre.

En ocasiones puede ser de gran utilidad combinar una gráfica de series de tiempo con otra de las representaciones gráficas que se han considerado anteriormente. J. Stuart Hunter (*The American Statistician,* vol. 42, p. 54) ha sugerido combinar el diagrama de tallo y hoja con una gráfica de series de tiempo para formar un **diagrama punto-dígito** (***digidot***).

En la figura 2-14 se muestra un diagrama punto-dígito para las observaciones de la resistencia a la compresión de la tabla 2-1, bajo el supuesto de que estas observaciones se registraron en el orden en que ocurrieron. Esta gráfica presenta de manera eficaz la variabilidad global de los datos de la resistencia a la compresión y muestra al mismo tiempo la variabilidad de las mediciones en el tiempo. La impresión general es que la resistencia a la compresión varía alrededor del valor medio de 162.67, pero no hay ningún patrón definido evidente de esta variabilidad en el tiempo.

El diagrama punto-dígito de la figura 2-15 cuenta una historia diferente. Esta gráfica resume 30 observaciones de la concentración del producto final de un proceso químico, donde las observaciones se registran en intervalos de tiempo de una hora. Esta gráfica indica que en las 20 primeras horas de operación del proceso se produjeron concentraciones generalmente arriba de 85 g/l, pero que después de la muestra 20 quizás haya ocurrido algo en el proceso que dio como resultado concentraciones más bajas. Si esta variabilidad en la concentración del producto final puede reducirse, entonces es posible mejorar la operación de este proceso.

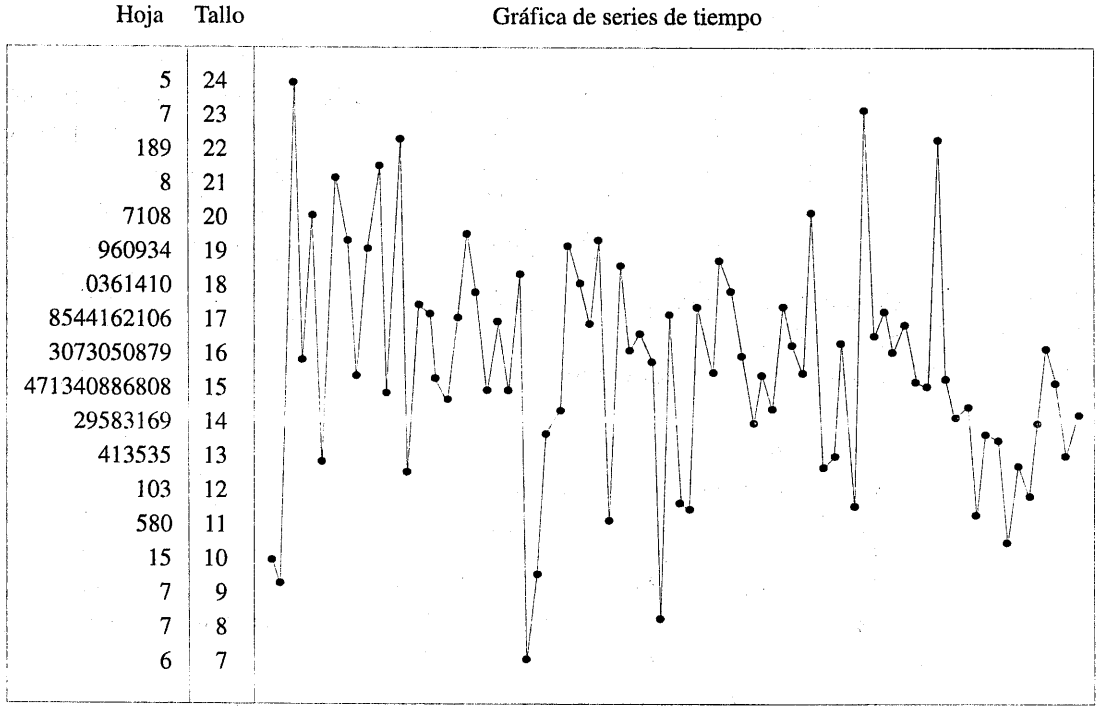


Figura 2-14 Diagrama punto-dígito de los datos de la resistencia a la compresión de la tabla 2-1.

Hoja	Tallo
8	9e
6	9s
45	9f
2333	9t
0010000	9z
99998	8e
66677	8s
45	8f
23	8t
1	8z

Gráfica de series de tiempo

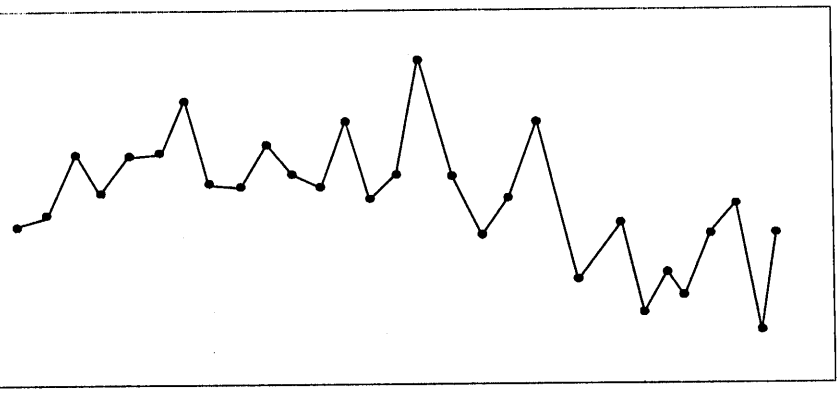


Figura 2-15 Diagrama punto-dígito de las lecturas de la concentración de un proceso químico, observadas cada hora.

EJERCICIOS DE LA SECCIÓN 2-5

2-21. El Colegio de Ingeniería y Ciencia Aplicada de la Universidad Estatal de Arizona tenía un sistema de computación VAX. Se registraron los tiempos de respuesta para 20 tareas consecutivas, los cuales se muestran a continuación en orden (léalos hacia abajo, y después de izquierda a derecha).

5.3	6.2	8.5	12.4
5.0	5.9	4.7	3.9
9.5	7.2	11.2	8.1
10.1	10.0	7.3	9.2
5.8	12.2	6.4	10.5

Construya e interprete una gráfica de series de tiempo de estos datos.

2-22. Los datos siguientes corresponden a las mediciones de la viscosidad de un producto químico observadas cada hora (léalos hacia abajo, y después de izquierda a derecha).

47.9	48.8	48.6	43.2	43.0
47.9	48.1	48.0	43.0	42.8
48.6	48.3	47.9	43.5	43.1
48.0	47.2	48.3	43.1	43.2
48.4	48.9	48.5	43.0	43.6
48.1	48.6	48.1	42.9	43.2
48.0	48.0	48.0	43.6	43.5
48.6	47.5	48.3	43.3	43.0

a) Construya e interprete un diagrama punto-dígito o bien un diagrama de tallo y hoja y, aparte, una gráfica de series de tiempo de estos datos.

b) Las especificaciones para la viscosidad del producto son 48 ± 2. ¿Qué conclusiones pueden sacarse del desempeño del proceso?

2-23. En una prueba de laboratorio se mide la fuerza de desconexión de un conector. A continuación se muestran los datos de 40 ejemplares de prueba (en el orden de medición, léalas hacia abajo, y después de izquierda a derecha).

241	220	249	209
258	194	251	212
237	245	238	185
210	209	210	187
194	201	198	218
225	195	199	190
248	255	183	175
203	245	213	178
195	235	236	175
249	220	245	190

a) Construya una gráfica de series de tiempo de los datos.

b) Construya e interprete un diagrama punto-dígito o bien un diagrama de tallo y hoja de los datos.

2-24. En su libro *Time Series Analysis, Forecasting, and Control,* Prentice-Hall, G.E.P. Box, G.M. Jenkins y G.C. Reinsel presentan las lecturas de la concentración en un proceso químico tomadas cada dos horas. Parte de estos datos se muestran a continuación (léalos hacia abajo y después de izquierda a derecha).

17.0	16.7	17.1	17.5	17.6
16.6	174.	17.4	18.1	17.5
16.3	17.2	17.4	17.5	16.5
16.1	17.4	17.5	17.4	17.8
17.1	17.4	17.4	17.4	17.3
16.9	17.0	17.6	17.1	17.3
16.8	17.3	17.4	17.6	17.1
17.4	17.2	17.3	17.7	17.4
17.1	17.4	17.0	17.4	16.9
17.0	16.8	17.8	17.8	17.3

Construya e interprete un diagrama punto-dígito o bien un diagrama de tallo y hoja de estos datos.

2-25. En la tabla 2-4 se presentan los números de manchas solares Wolfer anuales de 1770 a 1869. (Para un interesante análisis e interpretación de estos números, véase el libro de Box, Jenkins y Reinsel citado en el ejercicio 2-24. Este análisis requiere ciertos conocimientos avanzados de estadística y construcción de modelos estadísticos.)

a) Construya una gráfica de series de tiempo de estos datos.

b) Construya e interprete un diagrama punto-dígito o bien un diagrama de tallo y hoja de estos datos.

Tabla 2-4 Número de manchas solares anuales

1770	101	1795	21	1820	16	1845	40
1771	82	1796	16	1821	7	1846	62
1772	66	1797	6	1822	4	1847	98
1773	35	1798	4	1823	2	1848	124
1774	31	1799	7	1824	8	1849	96
1775	7	1800	14	1825	17	1850	66
1776	20	1801	34	1826	36	1851	64
1777	92	1802	45	1827	50	1852	54
1778	154	1803	43	1828	62	1853	39
1779	125	1804	48	1829	67	1854	21
1780	85	1805	42	1830	71	1855	7
1781	68	1806	28	1831	48	1856	4
1782	38	1807	10	1832	28	1857	23
1783	23	1808	8	1833	8	1858	55
1784	10	1809	2	1834	13	1859	94
1785	24	1810	0	1835	57	1860	96
1786	83	1811	1	1836	122	1861	77
1787	132	1812	5	1837	138	1862	59
1788	131	1813	12	1838	103	1863	44
1789	118	1814	14	1839	86	1864	47
1790	90	1815	35	1840	63	1865	30
1791	67	1816	46	1841	37	1866	16
1792	60	1817	41	1842	24	1867	7
1793	47	1818	30	1843	11	1868	37
1794	41	1819	24	1844	15	1869	74

Tabla 2-5 Millas de vuelo de aerolíneas de pasajeros en el Reino Unido

Mes	1964	1965	1966	1967	1968	1969	1970
Enero	7.269	8.350	8.186	8.334	8.639	9.491	10.840
Febrero	6.775	7.829	7.444	7.899	8.772	8.919	10.436
Marzo	7.819	8.829	8.484	9.994	10.894	11.607	13.589
Abril	8.371	9.948	9.864	10.078	10.455	8.852	13.402
Mayo	9.069	10.638	10.252	10.801	11.179	12.537	13.103
Junio	10.248	11.253	12.282	12.953	10.588	14.759	14.933
Julio	11.030	11.424	11.637	12.222	10.794	13.667	14.147
Agosto	10.882	11.391	11.577	12.246	12.770	13.731	14.057
Septiembre	10.333	10.665	12.417	13.281	13.812	15.110	16.234
Octubre	9.109	9.396	9.637	10.366	10.857	12.185	12.389
Noviembre	7.685	7.775	8.094	8.730	9.290	10.645	11.594
Diciembre	7.682	7.933	9.280	9.614	10.925	12.161	12.772

2-26. En su libro *Forecasting and Time Series Analysis*, 2a. edición (McGraw-Hill) D. C. Montgomery, L.A. Johnson y J.S. Gardiner analizan los datos de la tabla 2-5, que corresponden a las millas totales de vuelo de aerolíneas de pasajeros en el Reino Unido, 1964-1970 (en millones de millas).

a) Construye una gráfica de series de tiempo de los datos y comente las características de los datos que sean evidentes.

b) Construya e interprete un diagrama punto-dígito o bien un diagrama de tallo y hoja de estos datos.

Ejercicios complementarios

2-27. Un operador mide seis veces la concentración de una solución utilizando el mismo instrumento. Obtiene los siguientes datos: 63.2, 67.1, 65.8, 64.0, 65.1 y 65.3 (g/l).

a) Calcule la media muestral. Suponga que el valor deseable especificado para esta solución es 65.0 g/l. ¿Considera el lector que el valor de la media muestral calculado aquí está lo suficientemente cerca del valor especificado para dar el visto bueno a la solución por cumplir con las especificaciones? Explique el razonamiento.

b) Calcule la varianza muestral y la desviación estándar muestral.

c) Suponga que al medir la concentración, el operador necesita calibrar un aparato y usar un material reactivo. ¿Cuáles serían las principales fuentes de variabilidad en este experimento? ¿Por qué es deseable tener una varianza pequeña de estas mediciones?

2-28. Una muestra de seis resistores produjo las siguientes resistencias (ohms): $x_1 = 45$, $x_2 = 38$, $x_3 = 47$, $x_4 = 41$, $x_5 = 35$ y $x_6 = 43$.

a) Calcule la varianza muestral y la desviación estándar muestral empleando el método de la ecuación 1-4.

b) Calcule la varianza muestral y la desviación estándar muestral empleando la definición de la ecuación 1-3. Explique por qué los resultados obtenidos con la ecuación 1-4 son iguales.

c) Reste 35 a cada una de las mediciones de la resistencia originales y calcule s^2 y s. Compare los resultados con los obtenidos en los incisos a) y b) y explique los descubrimientos encontrados.

d) Si las resistencias fueran 450, 380, 470, 410, 350 y 430 ohms, ¿podrían usarse los resultados de los incisos anteriores de este problema para encontrar s^2 y s?

2-29. Un artículo en *Transactions of the Institution of Chemical Engineers* (vol. 34, pp. 280-293) reportó datos de un experimento que inves-

tigaba el efecto de distintas variables del proceso sobre la oxidación de la fase de vapor de la naftalina. Una muestra de la conversión porcentual en moles de la naftalina en anhídrido maleico es: 4.2, 4.7, 4.7, 5.0, 3.8, 3.6, 3.0, 5.1, 3.1, 3.8, 4.8, 4.0, 5.2, 4.3, 2.8, 2.0, 2.8, 3.3, 4.8, 5.0.

a) Calcule el rango muestral, la varianza muestral y la desviación estándar.

b) Calcule de nuevo el rango muestral, la varianza muestral y la desviación estándar, pero restando antes 1.0 de cada observación. Compare los resultados con los obtenidos en el inciso a). ¿Hay algo "especial" en cuanto a la constante 1.0, o cualquier otro valor elegido arbitrariamente hubiera producido los mismos resultados?

2-30. Considere las dos muestras presentadas a continuación:

Muestra 1: 10, 9, 8, 7, 8, 6, 10, 6
Muestra 2: 10, 6, 10, 6, 8, 10, 8, 6

a) Calcule el rango muestral de ambas muestras. ¿Podría concluirse que las dos presentan la misma variabilidad? Explique su respuesta.

b) Calcule la desviación estándar muestral de las dos muestras. ¿Estas cantidades indican que ambas muestras tienen la misma variabilidad? Explique su respuesta.

c) Escriba un enunciado breve contrastando el rango muestral y la desviación estándar muestral como medidas de variabilidad.

2-31. Un artículo en *Quality Engineering* (vol. 4, pp. 487-495) presenta datos de viscosidad de un proceso químico. Una muestra de estos datos se presenta a continuación.

13.3	14.9	15.8	16.0
14.5	13.7	13.7	14.9
15.3	15.2	15.1	13.6
15.3	14.5	13.4	15.3
14.3	15.3	14.1	14.3
14.8	15.6	14.8	15.6

(*continuación*)

15.2	15.8	14.3	16.1
14.5	13.3	14.3	13.9
14.6	14.1	16.4	15.2
14.1	15.4	16.9	14.4
14.3	15.2	14.2	14.0
16.1	15.2	16.9	14.4
13.1	15.9	14.9	13.7
15.5	16.5	15.2	13.8
12.6	14.8	14.4	15.6
14.6	15.1	15.2	14.5
14.3	17.0	14.6	12.8
15.4	14.9	16.4	16.1
15.2	14.8	14.2	16.6
16.8	14.0	15.7	15.6

a) Leyendo hacia abajo y de izquierda a derecha, haga una gráfica de series de tiempo de todos los datos y comente cualquier característica de los datos que salga a relucir en esta gráfica.

b) Considere la idea de que las 40 primeras observaciones se generaron de un proceso específico, en tanto que las 40 restantes se generaron de un proceso diferente. ¿El diagrama indica que los dos procesos producen resultados similares?

c) Calcule la media muestral y la varianza muestral de las 40 primeras observaciones; después calcule estos valores para las 40 últimas observaciones. ¿Estas cantidades indican que ambos procesos produducen el mismo nivel medio? ¿La misma variabilidad? Explique sus respuestas.

2-32. Los datos presentados en la tabla 2-6 corresponden a las ventas mensuales de champaña en Francia (1962-1969) en miles de botellas.

a) Construya una gráfica de series de tiempo de los datos y comente cualquier característica de los datos que salga a relucir en esta gráfica.

b) Especule sobre la forma en que se usaría un procedimiento gráfico para pronosticar las ventas mensuales de champaña para el año 1970.

Tabla 2-6 Ventas de champaña en Francia

Año	Mes	Ventas	Año	Mes	Ventas	Año	Mes	Ventas
1962	Ene.	2.851	1963	Ene.	2.541	1964	Ene.	3.113
	Feb.	2.672		Feb.	2.475		Feb.	3.006
	Mar.	2.755		Mar.	3.031		Mar.	4.047
	Abr.	2.721		Abr.	3.266		Abr.	3.523
	May.	2.946		May.	3.776		May.	3.937
	Jun.	3.036		Jun.	3.230		Jun.	3.986
	Jul.	2.282		Jul.	3.028		Jul.	3.260
	Ago.	2.212		Ago.	1.759		Ago.	1.573
	Sept.	2.922		Sept.	3.595		Sept.	3.528
	Oct.	4.301		Oct.	4.474		Oct.	5.211
	Nov.	5.764		Nov.	6.838		Nov.	7.614
	Dic.	7.132		Dic.	8.357		Dic.	9.254
1965	Ene.	5.375	1966	Ene.	3.633	1967	Ene.	4.016
	Feb.	3.088		Feb.	4.292		Feb.	3.957
	Mar.	3.718		Mar.	4.154		Mar.	4.510
	Abr.	4.514		Abr.	4.121		Abr.	4.276
	May.	4.520		May.	4.647		May.	4.968
	Jun.	4.539		Jun.	4.753		Jun.	4.677
	Jul.	3.663		Jul.	3.965		Jul.	3.523
	Ago.	1.643		Ago.	1.723		Ago.	1.821
	Sept.	4.739		Sept.	5.048		Sept.	5.222
	Oct.	5.428		Oct.	6.922		Oct.	6.873
	Nov.	8.314		Nov.	9.858		Nov.	10.803
	Dic.	10.651		Dic.	11.331		Dic.	13.916
1968	Ene.	2.639	1969	Ene.	3.934			
	Feb.	2.899		Feb.	3.162			
	Mar.	3.370		Mar.	4.286			
	Abr.	3.740		Abr.	4.676			
	May.	2.927		May.	5.010			
	Jun.	3.986		Jun.	4.874			
	Jul.	4.217		Jul.	4.633			
	Ago.	1.738		Ago.	1.659			
	Sept.	5.221		Sept.	5.591			
	Oct.	6.424		Oct.	6.981			
	Nov.	9.842		Nov.	9.851			
	Dic.	13.076		Dic.	12.670			

2-33. Un fabricante de resortes espirales está interesado en implementar un sistema de control de calidad para monitorear su proceso de producción. Como parte de este sistema de calidad, se decide registrar el número de resortes espirales que no se ajustan a las especificaciones en cada lote de producción de 50 unidades. Durante 40 días de producción, se recabaron datos de 40 lotes, los cuales se presentan aquí.

a) Construya un diagrama de tallo y hoja de los datos.

b) Encuentre el promedio y la desviación estándar muestral.

c) Construya una gráfica de series de tiempo de los datos. ¿Hay evidencia de que hubiera un aumento o una disminución en el número promedio de resortes espirales que no se ajustan a las especificaciones, fabricados durante los 40 días? Explique su respuesta. Lea los datos horizontalmente.

9	12	6	9	7	14	12	4	6	7
8	5	9	7	8	11	3	6	7	7
11	4	4	8	7	5	6	4	5	8
19	19	18	12	11	17	15	17	13	13

2-34. Se hace el monitoreo de un canal de comunicaciones registrando el número de errores en una cadena de 1 000 bits. Se presentan a continuación los datos de 20 de estas cadenas.

a) Construya un diagrama de tallo y hoja de los datos.

b) Encuentre el promedio y la desviación estándar muestral.

c) Construya una gráfica de series de tiempo de los datos. ¿Hay evidencia de que hubiera un aumento o una disminución en el número de errores en una cadena? Explique su respuesta. Lea los datos horizontalmente.

3	1	0	1	3	2	4	1	3	1
1	1	2	3	3	2	0	2	0	1

2-35. Transformaciones. En algunos conjuntos de datos, una transformación mediante la aplicación de una función matemática a los datos originales, como $\sqrt{y}$ o $\log y$, puede dar como resultado datos con los que es más sencillo trabajar estadísticamente que con los datos originales. Para ilustrar el efecto de una transformación, considérense los siguientes datos, que representan los ciclos para producir el desgaste de un producto textil: 675, 3 650, 175, 1 150, 290, 2 000, 100, 375.

a) Construya una gráfica de caja y comente la forma de la distribución de los datos.

b) Transforme los datos usando logaritmos; es decir, y^* (nuevo valor) = $\log y$ (valor antiguo). Construya una gráfica de caja de los datos transformados y comente el efecto de la transformación.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

2-36. Un experimento para investigar el tiempo de duración en horas de un componente electrónico consiste en colocar las partes en una celda de prueba y tenerlas en operación durante 100 horas bajo condiciones de temperatura elevada. (A ésta se le llama una prueba de duración "acelerada".) Se probaron ocho componentes con los siguientes tiempos de falla resultantes:

$$75, 63, 100^+, 36, 51, 45, 80, 90$$

La observación "100^+" indica que la unidad aún seguía funcionando a las 100 horas. ¿Existe alguna medida de localización significativa que pueda calcularse para estos datos? ¿Cuál es su valor numérico?

2-37. Suponga que se tiene una muestra $x_1, x_2, \ldots, x_n$ y que se ha calculado $\bar{x}_n$ y s_n^2 para la muestra. Entonces se dispone de una observación $(n + 1)$. Sean $\bar{x}_{n+1}$ y s_{n+1}^2 la media muestral y la varianza muestral de la muestra utilizando las $n + 1$ observaciones.

a) Indique cómo puede calcularse $\bar{x}_{n+1}$ utilizando $\bar{x}_n$ y x_{n+1}.

b) Demuestre que

$$ns_{n+1}^2 = (n-1)\, s_n^2 + \frac{n}{n+1}(x_{n+1} - \bar{x}_n)^2.$$

c) Use los resultados de los incisos a) y b) para calcular el nuevo promedio muestral y la nueva desviación estándar muestral para los datos del ejercicio 2-28, cuando la nueva observación es $x_7 = 46$.

2-38. La media recortada. Suponga que los datos están ordenados en orden creciente, que se omiten $T\%$ de las observaciones de ambos extremos y que se calcula la media muestral de los números restantes. A la cantidad resultante se le llama *media recortada.* La media recortada por lo general está entre la media muestral $\bar{x}$ y la mediana muestral $\tilde{x}$. ¿Por qué?

a) Calcule la media muestral recortada con el 10% para los datos de rendimiento del ejercicio 2-4.

b) Calcule la media muestral recortada con el 20% para los datos de rendimiento del ejercicio 2-4 y compárela con la cantidad encontrada en el inciso a).

c) Compare los valores calculados en los incisos a) y b) con la media y la mediana muestral para los datos del ejercicio 2-4. ¿Hay mucha diferencia en estas cantidades? ¿Por qué?

2-39. La media recortada. Suponga que el tamaño de la muestra n es tal que la cantidad $nT/100$ no es un entero. Desarrolle un procedimiento para obtener una media recortada en este caso.

3 Probabilidad

PLAN GENERAL DEL CAPÍTULO

3-1 ESPACIOS MUESTRALES Y EVENTOS

3-1.1 Introducción

3-1.2 Espacios muestrales

3-1.3 Eventos

3-2 INTERPRETACIÓN DE LA PROBABILIDAD

3-2.1 Introducción

3-2.2 Axiomas de probabilidad

3-3 REGLAS DE ADICIÓN

3-4 PROBABILIDAD CONDICIONAL

3-5 REGLAS DE MULTIPLICACIÓN Y DE PROBABILIDAD TOTAL

3-5.1 Regla de multiplicación

3-5.2 Regla de probabilidad total

3-6 INDEPENDENCIA

3-7 TEOREMA DE BAYES

3-8 VARIABLES ALEATORIAS

3-1 ESPACIOS MUESTRALES Y EVENTOS

3-1.1 Introducción

Cuando se mide la corriente eléctrica en un alambre delgado de cobre, se está realizando un **experimento**. Sin embargo, los resultados de las repeticiones de la medición día con día pueden diferir un tanto debido a ligeras variaciones en las variables del experimento que no están sujetas a control, incluyendo los cambios en la temperatura ambiente, las ligeras variaciones en el manómetro (presión atmosférica) y las pequeñas impurezas en la composición química del alambre si se seleccionan lugares diferentes, y las variaciones en la fuente de la corriente. Por consi-

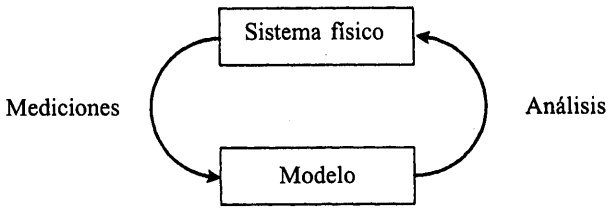


Figura 3-1 Iteración continua entre modelo y sistema físico.

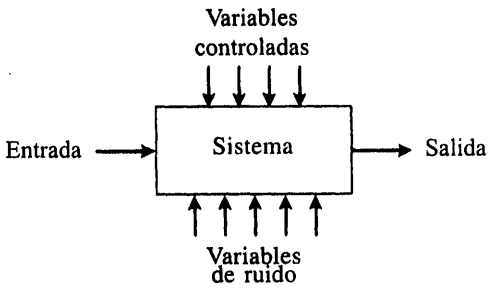


Figura 3-2 Las variables de ruido afectan la transformación de entradas en salidas.

guiente, se dice que este experimento (así como muchos que realizamos) tiene un componente **aleatorio**. En algunos casos, las variaciones aleatorias que ocurren son tan pequeñas, en comparación con los objetivos del experimento, que pueden ignorarse. Sin embargo, casi siempre está presente la variación, y su magnitud puede ser lo suficientemente grande para que las conclusiones importantes del experimento no sean evidentes. En estos casos, los métodos presentados en este libro para hacer un modelo y el análisis de los resultados experimentales son muy valiosos.

Sin importar qué tan cuidadosamente se diseñe y se conduzca un experimento, es frecuente que ocurran variaciones. Nuestra meta es comprender, cuantificar y hacer un modelo del tipo de variaciones que se encuentran con frecuencia. Cuando se incorpora la variación en nuestros razonamientos y análisis, es posible formular juicios fundamentados de los resultados que no se invalidan por la variación.

Los modelos y análisis que incluyen la variación no son diferentes de los modelos usados en otras áreas de la ingeniería y las ciencias. En la figura 3-1 se muestran los componentes importantes. Se desarrolla un modelo matemático (o abstracción) del sistema físico. No necesita ser una abstracción perfecta. Por ejemplo, las leyes de Newton no son descripciones perfectas de nuestro universo físico. Aun así, constituyen modelos útiles que pueden estudiarse y analizarse para cuantificar en forma aproximada el desempeño de un amplio rango de productos de la ingeniería. Dada una abstracción matemática que esté validada con mediciones de nuestro sistema, podemos usar el modelo para comprender, describir y cuantificar aspectos importantes del sistema físico, así como para predecir la respuesta del sistema con las entradas.

En este libro se exponen modelos que toman en cuenta las variaciones en las salidas de un sistema, aun cuando las variables sujetas a control no cambien de manera intencionada durante el estudio. En la figura 3-2 se muestra la representación gráfica de un modelo que incorpora entradas no controlables (ruido) que se combinan con las entradas controlables para producir la salida del sistema. Debido a las entradas no controlables, los mismos ajustes de las entradas controlables no resultan en salidas idénticas cada vez que se mide el sistema.

Para el ejemplo de la medición de la corriente eléctrica en un alambre de cobre, el modelo del sistema podría ser simplemente la ley de Ohm, como se señaló en la sección 1-3. Debido a las entradas no controlables, son de esperarse variaciones en las mediciones de la corriente. La ley de Ohm podría ser una aproximación adecuada. Sin embargo, si las variaciones son grandes respecto del uso previsto del dispositivo bajo estudio, quizá sería necesario ampliar el modelo a fin de incluir la variación. Muchas veces es difícil especular sobre la magnitud de las variaciones sin mediciones empíricas. Sin embargo, con suficientes mediciones es posible llegar a una aproximación de la magnitud de la variación y considerar su efecto sobre el desempeño de otros

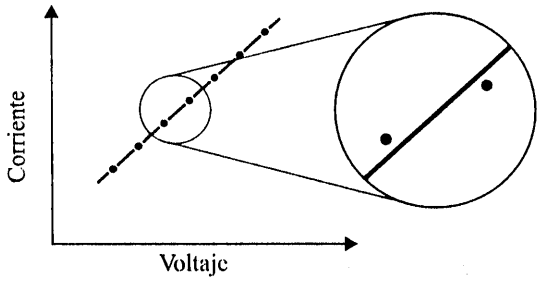


Figura 3-3 Un examen más detallado del sistema identifica desviaciones del modelo.

dispositivos, tales como amplificadores, en el circuito. Por lo tanto, estamos confirmando el modelo de la figura 3-2 como una descripción más conveniente de la medición de la corriente. En consecuencia, las técnicas que se presentan en este texto para el análisis de modelos que incluyen variación, con frecuencia resultan de utilidad (véase también la figura 3-3).

Como un ejemplo más, en el diseño de un sistema de comunicaciones, como una red de computadoras o una red de comunicación de voz, la capacidad de información disponible para atender a los usuarios que utilizan la red es una consideración de diseño importante. Para la comunicación de voz, es necesario adquirir suficientes líneas externas de la compañía telefónica para cumplir con los requerimientos de un negocio. Suponiendo que cada línea puede transmitir una sola conversación, ¿cuántas líneas deberán contratarse? Si se adquieren muy pocas líneas, las llamadas pueden retrasarse o perderse. La adquisición de demasiadas líneas incrementa los costos. De manera creciente, se necesita el diseño y desarrollo de productos para satisfacer los requerimientos del cliente *a un costo competitivo.*

En el diseño del sistema de comunicación de voz, se necesita un modelo para el número de llamadas y la duración de las mismas. Inclusive saber que en promedio las llamadas ocurren cada cinco minutos y que duran cinco minutos, no es suficiente. Si las llamadas llegaran precisamente en intervalos de cinco minutos y duraran exactamente cinco minutos, entonces bastaría una sola línea telefónica. Sin embargo, la más ligera variación en el número de llamadas o en su duración ocasionaría que algunas llamadas fueran bloqueadas por otras (véase la figura 3-4). Un sistema diseñado sin considerar la variación resultará funestamente inadecuado para uso práctico. Es necesario que el modelo para el número y duración de las llamadas incluya la variación

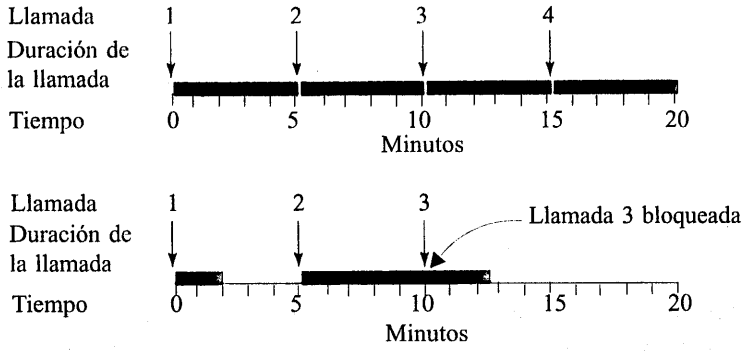


Figura 3-4 La variación provoca interrupciones en el sistema.

como componente integral. El análisis de modelos que incluyan la variación es importante para el diseño del sistema telefónico.

3-1.2 Espacios muestrales

Definición

Un **experimento aleatorio** es aquel que puede producir resultados diferentes, aun cuando se repita siempre de la misma manera.

La energía consumida en una reacción química puede variar cuando el experimento se repite en tiempos diferentes o en laboratorios diferentes. Se trata de un experimento aleatorio con varios resultados. Supóngase que algunos de los componentes electrónicos de un día de producción no cumplen con las especificaciones de regulación del encendido de un sistema de alto acabado. Al seleccionar dos piezas del lote producido ese día tomando arbitrariamente dos de ellas y observar si cada pieza cumple con las especificaciones de regulación del encendido puede considerarse un experimento aleatorio. Los resultados dependen de las dos piezas particulares que por casualidad se hayan seleccionado. Si este experimento se repitiera, podrían escogerse dos piezas diferentes. Por consiguiente, podrían obtenerse resultados de la regulación del encendido diferentes.

Para hacer el modelo de un experimento aleatorio y analizarlo, es necesario entender el conjunto de los resultados posibles del experimento. En la presente introducción a la probabilidad se hace uso de los conceptos básicos de conjuntos y de las operaciones con los mismos. Se supone que el lector está familiarizado con estos temas.

Definición

Al conjunto de todos los resultados posibles de un experimento aleatorio se le llama **espacio muestral** del experimento. El espacio muestral se denota por S.

Es frecuente definir el espacio muestral con base en los objetivos del análisis.

EJEMPLO 3-1

Considérese un experimento en el que se selecciona un conector y se mide su espesor. Los valores posibles del espesor dependen de la resolución del instrumento de medición, así como de los límites superior e inferior del espesor. Sin embargo, podría resultar conveniente definir el espacio muestral simplemente como la recta real R,

$$S = R$$

aun cuando no puede ocurrir un valor negativo del espesor.

Si el único objetivo del análisis es considerar si una pieza particular tiene espesor bajo, medio o alto, entonces el espacio muestral podría tomarse como el conjunto de los tres resultados

$$S = \{bajo, medio, alto\}$$

Si el único objetivo del análisis es considerar si una pieza particular cumple o no con las especificaciones de fabricación, entonces el espacio muestral podría simplificarse al conjunto de los dos resultados

$$S = \{sí, no\}$$

que indica si la pieza cumple o no con las especificaciones.

EJEMPLO 3-2

Si se seleccionan y miden dos conectores, entonces la extensión de la recta real R llevará el espacio muestral al plano

$$S = R \times R$$

Si el único objetivo del análisis es considerar si las piezas particulares cumplen o no con las especificaciones de fabricación, entonces cualquiera de las dos puede cumplir con ellas o no. Se abrevian *sí* y *no* como *s* y *n*. Si el par ordenado *sn* indica que el primer conector cumple con las especificaciones y el segundo no lo hace, entonces el espacio muestral puede representarse por los cuatro resultados

$$S = \{ss, sn, ns, nn\}$$

Si sólo nos interesáramos en el número de piezas de la muestra que cumple con las especificaciones, el espacio muestral podría resumirse como

$$S = \{0, 1, 2\}$$

Como otro ejemplo, considérese un experimento en el que el espesor se mide hasta que un conector no cumple con las especificaciones. El espacio muestral puede representarse como

$$S = \{n, sn, ssn, sssn, ssssn, \text{ y así sucesivamente}\}$$

En los experimentos aleatorios que implican seleccionar artículos de un lote, se indicará si el artículo seleccionado se reemplaza o no antes de seleccionar el siguiente. Por ejemplo, si el lote se compone de tres artículos $\{a, b, c\}$ y el experimento consiste en seleccionar dos artículos **sin reemplazo**, el espacio muestral puede representarse como $S = \{ab, ac, ba, bc, ca, cb\}$. Sin embargo, si los artículos se reemplazan antes de seleccionar el siguiente, se dice que el muestreo es **con reemplazo**. Entonces, los resultados posibles son $S = \{aa, ab, ac, ba, bb, bc, ca, cb, cc\}$.

3-1.3 Eventos

Con frecuencia nos interesamos en un conjunto de resultados relacionados de un experimento aleatorio.

Definición

Un **evento** es un subconjunto del espacio muestral de un experimento aleatorio.

Con frecuencia tenemos interés en describir nuevos eventos a partir de combinaciones de los eventos existentes. Dado que los eventos son subconjuntos, es posible usar las operaciones básicas con conjuntos tales como la unión, intersección y el complemento para formar otros eventos de interés. Algunas de las operaciones básicas con conjuntos se resumen a continuación en términos de eventos:

La **unión** de dos eventos es el evento que consta de todos los resultados que están contenidos en cualquiera de los dos eventos. La unión se denota por $E_1 \cup E_2$.

La **intersección** de dos eventos es el evento que consta de todos los resultados que están contenidos en los dos eventos. La intersección se denota por $E_1 \cap E_2$.

El **complemento** de un evento en un espacio muestral es el conjunto de resultados en el espacio muestral que no están en el evento. El complemento del evento E se denota por E'.

EJEMPLO 3-3

En el ejemplo 3-2, suponga que el conjunto de todos los resultados para los que al menos una pieza cumple con las especificaciones se denota por E_1. Entonces,

$$E_1 = \{ss, sn, ns\}$$

El evento de que ninguna de las dos piezas cumpla con las especificaciones, denotado por E_2, sólo contiene el resultado, $E_2 = \{nn\}$. Otros ejemplos de eventos son $E_3 = \varnothing$, el conjunto vacío, y $E_4 = S$, el espacio muestral.

Si $E_5 = \{sn, ns, nn\}$, entonces

$$E_1 \cup E_5 = S \qquad E_1 \cap E_5 = \{sn, ns\} \qquad E_1' = \{nn\}$$

EJEMPLO 3-4

Un modelo de las mediciones del tiempo necesario para completar una reacción química podría ser el espacio muestral $S = (0, \infty)$, el conjunto de los números reales positivos. Sea

$$E_1 = \{x \mid 1 \leq x < 10\}$$

y sea

$$E_2 = \{x \mid 3 < x < 118\}$$

Entonces,

$$E_1 \cup E_2 = \{x \mid 1 \leq x < 118\} \quad \text{y} \quad E_1 \cap E_2 = \{x \mid 3 < x < 10\}$$

Además,

$$E_1' = \{x \mid x \geq 10\}$$

y

$$E_1' \cup E_2 = \{x \mid 10 \leq x < 118\}$$

EJEMPLO 3-5

Se analizan muestras de plástico de policarbonato para determinar la resistencia a las rayaduras y la resistencia a los impactos. Los resultados de 49 muestras se resumen a continuación:

		resistencia a las rayaduras	
		alta	baja
resistencia a los impactos	alta	40	4
	baja	2	3

Sea que A denote el evento de que una muestra tiene alta resistencia a los impactos y sea que B denote el evento de que una muestra tiene alta resistencia a las rayaduras. Determine el número de muestras en $A \cap B$, A' y $A \cup B$.

El evento $A \cap B$ consta de las 40 muestras para las que la resistencia a las rayaduras y a los impactos son altas. El evento A' consta de las 7 muestras para las que la resistencia a los impactos es baja. El evento $A \cup B$ consta de las 46 muestras para las que la resistencia a los impactos, la resistencia a las rayaduras o ambas son altas.

Con frecuencia se usan diagramas para ilustrar las relaciones entre conjuntos y estos diagramas también se usan para describir las relaciones entre eventos. Pueden usarse **diagramas de Venn** para representar un espacio muestral y los eventos en un espacio muestral. Por ejemplo, en la figura 3-5*a* el espacio muestral del experimento aleatorio se representa como los puntos del rectángulo S. Los eventos A y B son los subconjuntos de puntos en las regiones

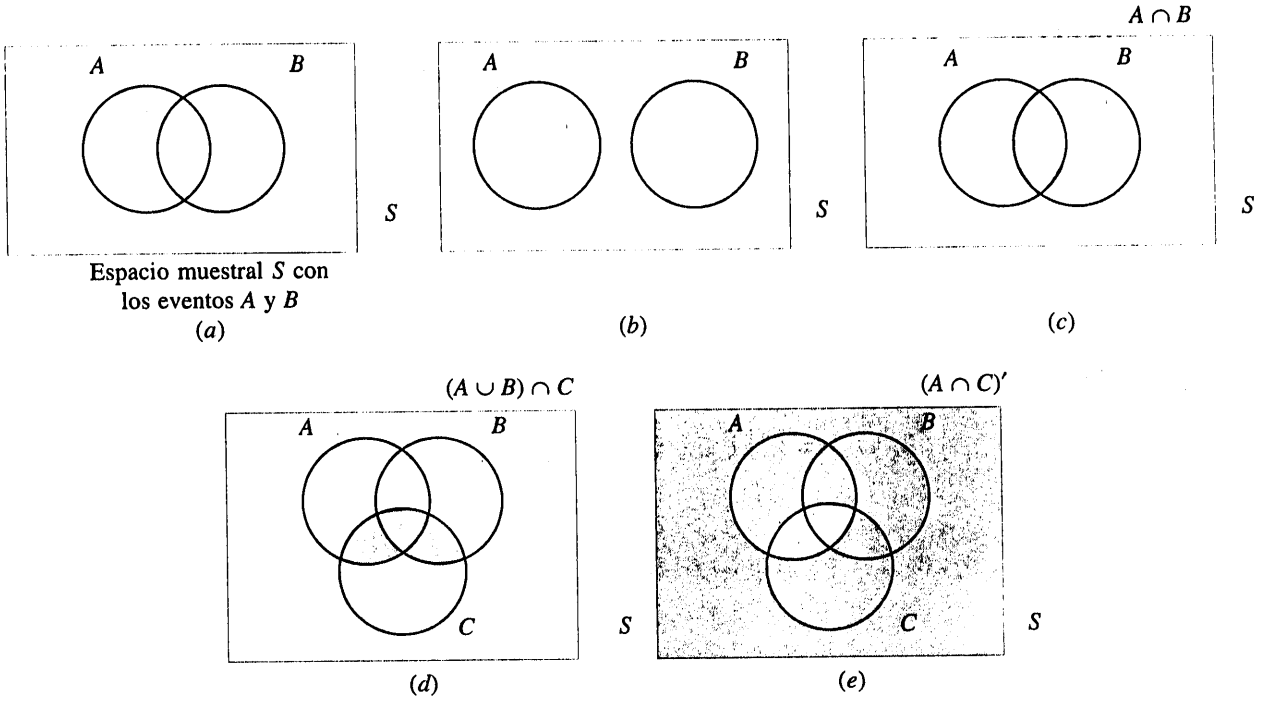


Figura 3-5 Diagramas de Venn.

indicadas. En la figura 3-5*b* se ilustran dos eventos que no tienen resultados comunes; en las figuras 3-5*c* a 3-5*e* se ilustran otros eventos conjuntos.

Los espacios muestrales también pueden describirse gráficamente con **diagramas de árbol**. Cuando un espacio muestral puede construirse en varios pasos o etapas, cada una de las n_1 formas de completar el primer paso puede representarse como la rama de un árbol. Cada una de las formas de completar el segundo paso puede representarse como n_2 ramas que empiezan en los extremos de las ramas originales, y así sucesivamente.

EJEMPLO 3-6

Cada mensaje en un sistema de comunicación digital se clasifica de acuerdo a si se recibe dentro del tiempo especificado por el diseño del sistema. Si se clasifican tres mensajes, use un diagrama de árbol para representar el espacio muestral de los resultados posibles.

Cada mensaje puede recibirse a tiempo o retrasado. Los resultados posibles para los tres mensajes pueden representarse con ocho ramas en el diagrama de árbol mostrado en la figura 3-6.

EJEMPLO 3-7

Un fabricante de automóviles ofrece vehículos equipados con accesorios opcionales. El pedido de cada vehículo se hace

Con o sin transmisión automática.

Con o sin aire acondicionado.

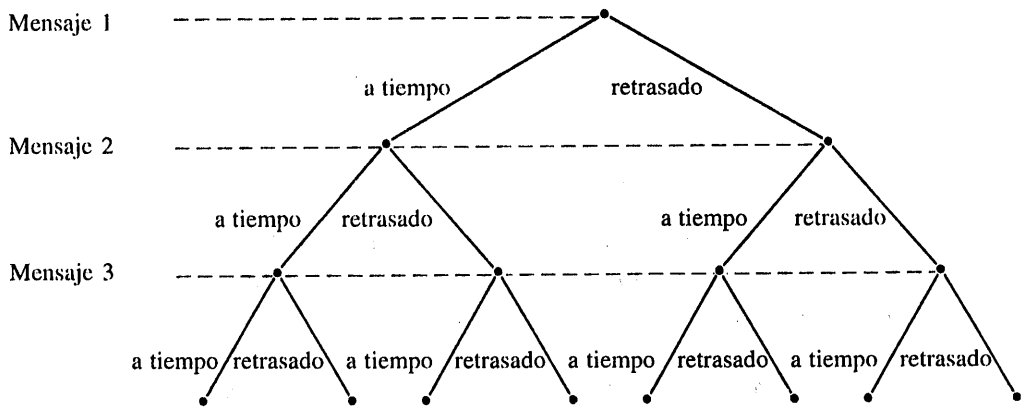


Figura 3-6 Diagrama de árbol para los tres mensajes.

Con una de tres opciones de un sistema estéreo.

Con uno de cuatro colores exteriores.

Si el espacio muestral consta del conjunto de todos los tipos de vehículos posibles, ¿cuál es el número de resultados en el espacio muestral? El espacio muestral contiene 48 resultados. El diagrama de árbol de los diferentes tipos de vehículos se muestra en la figura 3-7.

EJEMPLO 3-8

Considérese una ampliación del ejemplo anterior en la que el fabricante de automóviles ofrece otra opción: el color interior. Hay cuatro opciones para el color interior: rojo, negro, azul o café. Sin embargo,

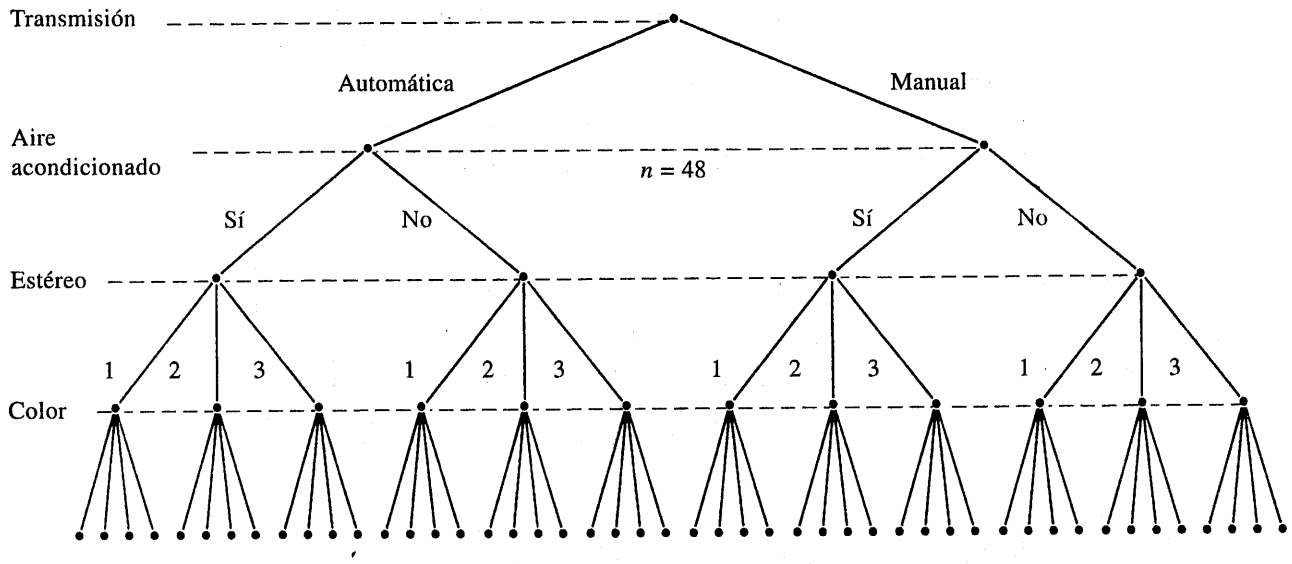


Figura 3-7 Diagrama de árbol para los diferentes tipos de vehículos.

Con un exterior rojo, sólo puede escogerse un interior negro o rojo.

Con un exterior blanco, puede escogerse cualquier color interior.

Con un exterior azul, sólo puede escogerse un interior negro, rojo o azul.

Con un exterior café, sólo puede escogerse un interior café.

Hay 12 tipos de vehículos con cada color exterior, pero el número de opciones para el color interior depende del color exterior. Como se muestra en la figura 3-8, es posible ampliar el diagrama de árbol para mostrar que hay 120 tipos diferentes de vehículos en el espacio muestral.

Dos eventos sin resultados en común presentan una importante relación.

Definición

Se dice que dos eventos, denotados como E_1 y E_2, tales que $E_1 \cap E_2 = \varnothing$ son **mutuamente excluyentes.**

Los dos eventos de la figura 3-5*b* son mutuamente excluyentes, en tanto que los dos eventos de la figura 3-5*a* no lo son.
A continuación se resumen otros resultados útiles en los que intervienen eventos. La definición del complemento de un evento implica que

$$(E')' = E$$

La ley distributiva para operaciones con conjuntos implica que

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

y

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

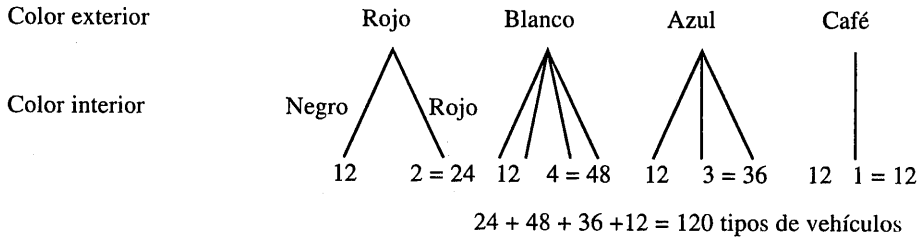


Figura 3-8 Diagrama de árbol para los diferentes tipos de vehículos con colores interiores.

Las leyes de DeMorgan implican que

$$(A \cup B)' = A' \cap B' \text{ y } (A \cap B)' = A' \cup B'$$

Asimismo, recuérdese que

$$A \cap B = B \cap A \text{ y } A \cup B = B \cup A$$

EJERCICIOS DE LA SECCIÓN 3-1

Presente una descripción razonable del espacio muestral para cada uno de los experimentos aleatorios de los ejercicios 3-1 al 3-15. Puede haber más de una interpretación aceptable para cada experimento. Describa cualquier supuesto que haga.

3-1. Cada una de tres piezas maquinadas se clasifica como arriba o abajo de las especificaciones de la pieza.

3-2. Cada uno de cuatro bits transmitidos se clasifica como con error o sin error.

3-3. En la inspección final de fuentes de poder electrónicas podrían ocurrir tres tipos de disconformidades: funcionales, secundarias y de acabado. Las fuentes de poder defectuosas se clasifican además según sea el tipo de disconformidad.

3-4. En la fabricación de cinta para grabación digital se usa una prueba electrónica para registrar el número de bits con error en un carrete de 350 pies.

3-5. En la fabricación de cinta para grabación digital, cada una de 24 pistas se clasifica de acuerdo a si contiene o no uno o más bits con error.

3-6. Un amperímetro con pantalla de tres dígitos se usa para medir corriente eléctrica en miliamperes.

3-7. Una báscula cuya pantalla muestra dos cifras decimales se usa para medir la alimentación de materiales en una planta química en toneladas.

3-8. Las dos preguntas siguientes aparecen en el cuestionario de una encuesta para empleados. Cada respuesta se elige de la escala de cinco puntos 1 (nunca), 2, 3, 4, 5 (siempre).
"¿La corporación tiene disposición para escuchar y evaluar imparcialmente las ideas nuevas?"
"¿Con qué frecuencia mis compañeros de trabajo son importantes en mi desempeño laboral general?"

3-9. Las oquedades en una placa de ferrita se clasifican como pequeñas, medianas o grandes. El número de oquedades en cada categoría se miden por una inspección visual de una muestra.

3-10. En un proceso de fabricación pueden producirse algunas piezas que no son aceptables. Cada una de tres partes se clasifica como aceptable o no aceptable.

3-11. En el pedido de un automóvil puede especificarse transmisión automática o estándar, con o sin aire acondicionado y cualquiera de los cuatro colores rojo, azul, negro o blanco. Describa el conjunto de los pedidos posibles para este experimento.

3-12. Una pieza moldeada por inyección muestreada pudo haberse hecho en cualquiera de dos prensas y en cualquiera de las ocho cavidades de cada prensa.

3-13. En el pedido de una computadora puede especificarse memoria de 4, 8 o 12 megabytes y capacidad de almacenamiento en disco duro de 200, 300 o 400 megabytes. Describa el conjunto de pedidos posibles.

3-14. Se hacen llamadas repetidas a una línea telefónica ocupada hasta que se consigue la conexión.

3-15. En un dispositivo de almacenamiento magnético se hacen tres intentos para leer datos antes de que se aplique un procedimiento de recuperación de errores que reposiciona la cabeza magnética. El procedimiento de recuperación de errores intenta tres reposicionamientos antes de enviar un mensaje de "aborto" al operador. Sea que

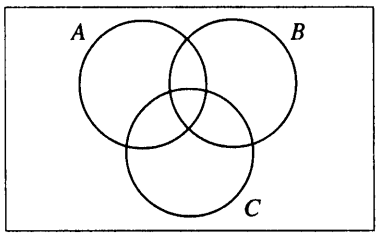


Figura 3-9 Diagrama de Venn para el ejercicio 3-16.

s denote el éxito de una operación de lectura,

f denote el fracaso de una operación de lectura,

F denote el fracaso del procedimiento de recuperación de errores,

S denote el éxito del procedimiento de recuperación de errores, y

A denote un mensaje de aborto enviado al operador.

Describa el espacio muestral de este experimento con un diagrama de árbol.

3-16. En el diagrama de Venn de la figura 3-9 se muestran tres eventos. Copie la figura y sombree la región que corresponda a cada uno de los eventos siguientes:

a) A'
b) $A \cap B$
c) $(A \cap B) \cup C$
d) $(B \cup C)'$
e) $(A \cap B)' \cup C$

3-17. En el diagrama de Venn de la figura 3-10 se muestran tres eventos. Copie la figura y sombree la región que corresponda a cada uno de los eventos siguientes:

a) A'
b) $(A \cap B) \cup (A \cap B')$

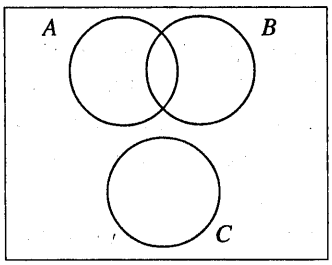


Figura 3-10 Diagrama de Venn para el ejercicio 3-17.

c) $(A \cap B) \cup C$
d) $(B \cup C)'$
e) $(A \cap B)' \cup C$

3-18. Se usa una báscula digital que da los pesos hasta el gramo más próximo.

a) ¿Cuál es el espacio muestral de este experimento?

Sea que

A denote el evento de que un peso exceda los 11 gramos, sea que B denote el evento de que un peso sea menor o igual que 15 gramos y sea que C denote el evento de que un peso sea mayor o igual que 8 gramos y menor que 12 gramos.

Describa los siguientes eventos:

b) $A \cup B$
c) $A \cap B$
d) A'
e) $A \cup B \cup C$
f) $(A \cup C)'$
g) $A \cap B \cap C$
h) $B' \cap C$
i) $A \cup (B \cap C)$

3-19. En una operación de moldeo por inyección se evalúan varias características de cada pieza producida. Sea que

A denote el evento de que una pieza cumple con los requerimientos de contracción del cliente,

B denote el evento de que una pieza cumple con los requerimientos de color del cliente, y

C denote el evento de que una longitud crítica cumple con los requerimientos del cliente.

a) Construya un diagrama de Venn que incluya estos eventos e indique la región del diagrama en la que una pieza cumple con todos los requerimientos del cliente. Sombree las áreas que representen lo siguiente:

b) $B \cap C$
c) $A' \cup B$
d) $A \cup B$
e) Si dos cualesquiera de estos eventos fueran mutuamente excluyentes, ¿qué tanto éxito tendría esta operación de producción?

3-20. Se transmiten cuatro bits en un canal de comunicación digital. Cada bit está distorsionado o se recibe sin distorsión. Sea que A_i denote el evento de que el i-ésimo bit esté distorsionado, $i = 1, ..., 4$.

a) Describa el espacio muestral para este experimento.

b) ¿Los eventos A_i' son mutuamente excluyentes?

c) Describa el resultado de cada uno de los eventos siguientes:

A_1

A_1'

$A_1 \cap A_2 \cap A_3 \cap A_4$

$(A_1 \cap A_2) \cup (A_3 \cap A_4)$

3-21. Se selecciona una muestra de tres calculadoras de una línea de fabricación y cada una de ellas se clasifica como defectuosa o aceptable. Sea que A, B y C denoten los eventos de que la primera, la segunda y la tercera calculadora esté defectuosa, respectivamente.

a) Describa el espacio muestral para este experimento con un diagrama de árbol.

Describa cada uno de los eventos siguientes:

b) A

c) B

d) $A \cap B$

e) $B \cup C$

3-22. Se analizan las placas plásticas circulares de policarbonato de un proveedor para la resistencia a las rayaduras y la resistencia a los impactos. Los resultados de 100 placas circulares se resumen a continuación.

		resistencia a las rayaduras	
		alta	baja
resistencia a	alta	80	9
los impactos	baja	6	5

Sea que A denote el evento de que una placa circular tenga alta resistencia a los impactos y sea que B denote el evento de que una placa circular tenga alta resistencia a las rayaduras. Determine el número de placas circulares en $A \cap B$, A' y $A \cup B$.

3-23. Muestras de una pieza de aluminio forjado se clasifican con base en el acabado de la superficie (en micropulgadas) y en las mediciones de la longitud. Los resultados de 100 piezas se resumen a continuación:

		longitud	
		excelente	buena
acabado de	excelente	75	7
la superficie	bueno	10	8

a) Sea que A denote el evento de que una muestra tiene un acabado de la superficie excelente y sea que B denote el evento de que una muestra tiene una longitud excelente. Determine el número de muestras en $A' \cap B$, B' y $A \cup B$.

b) Suponga que cada una de dos muestras debe clasificarse con base en el acabado de la superficie, excelente o bueno, y en la longitud, excelente o buena. Use un diagrama de árbol para representar los resultados posibles de este experimento.

3-24. Muestras de hule espuma de tres proveedores se clasifican de acuerdo a si cumplen o no con las especificaciones. Los resultados de 100 muestras se resumen a continuación.

		cumple	
		Sí	No
	1	18	2
proveedor	2	17	3
	3	50	10

Sea que A denote el evento de que una muestra es del proveedor 1 y sea que B denote el evento de que una muestra cumple con las especificaciones. Determine el número de muestras en $A' \cap B$, B' y $A \cup B$.

3-25. El tiempo de elevación de un reactor se mide en minutos (y fracciones de minutos). Sea el espacio muestral los números reales positivos. Defina los eventos A y B de la siguiente manera:

$$A = \{x \mid x < 72.5\}$$

y

$$B = \{x \mid x > 52.5\}$$

Describa cada uno de los eventos siguientes:

a) A'

b) B'

c) $A \cap B$
d) $A \cup B$

3-26. El tiempo de elevación de un reactor se mide en minutos (y fracciones de minutos). Sean los números reales positivos el espacio muestral para el tiempo de elevación de una carga. Considere los tiempos de elevación de dos cargas. Sea que A denote el evento de que el tiempo de elevación de ambas cargas sea menor que 72.5 minutos y sea que B denote el evento de que el tiempo de elevación de ambos procesamientos sea mayor que 52.5 minutos.

a) Describa gráficamente el espacio muestral para el tiempo de elevación de dos cargas. Indique cada uno de los eventos siguientes en la gráfica del espacio muestral:
b) A
c) B
d) A'
e) $A \cap B$
f) $A \cup B$

3-2 INTERPRETACIÓN DE LA PROBABILIDAD

3-2.1 Introducción

En este capítulo se introduce la probabilidad únicamente para los espacios muestrales con un conjunto finito (o infinito contable) de resultados. La restricción a estos espacios muestrales permite simplificar los conceptos, así como la presentación sin matemáticas avanzadas.

Definición

Un espacio muestral es **discreto** si contiene un conjunto finito (o contablemente infinito) de resultados.

Con frecuencia resulta conveniente cuantificar la veracidad, o posibilidad, de que ocurrirá un resultado de un experimento aleatorio. "Las posibilidades de que llueva hoy son de 30%" es un enunciado que cuantifica nuestro sentir en cuanto a la posibilidad de que llueva. La veracidad o factibilidad de un resultado se cuantifica asignando un número del intervalo [0, 1] al resultado (o un porcentaje de 0 a 100%). Los números más altos indican que el resultado es más factible que los números más bajos. Un cero indica que un resultado no ocurrirá. Un "uno" indica que un resultado ocurrirá con certeza.

La probabilidad de un resultado puede interpretarse como nuestra probabilidad subjetiva, o **grado de creencia**, de que el resultado ocurrirá. Diferentes individuos no dudarán en asignar probabilidades diferentes a los mismos resultados. Otra interpretación de la probabilidad se basa en el modelo conceptual de repetir en varias ocasiones el experimento aleatorio. La probabilidad de un resultado se interpreta como el valor límite de la proporción de veces que ocurre el resultado en n repeticiones del experimento aleatorio cuando n se incrementa sin restricciones. Por ejemplo, si se asigna la probabilidad 0.2 al resultado de que hay un pulsación viciada en una señal digital, podría interpretarse que esta asignación implicaría que, si se analizan muchas pulsaciones, aproximadamente 20% de las mismas estarán viciadas. Este ejemplo proporciona una interpretación de la probabilidad a través de la **frecuencia relativa**. La proporción, o fre-

cuencia relativa, de las repeticiones del experimento que producen el resultado es 0.2. Las probabilidades se eligen de tal forma que la suma de las probabilidades de todos los resultados de un experimento se acumulen hasta uno. Esta convención facilita la interpretación de la probabilidad como frecuencia relativa. En la figura 3-11 se ilustra el concepto de frecuencia relativa.

Las probabilidades de un experimento aleatorio suelen asignarse con base en un modelo razonable del sistema bajo estudio. Uno de los enfoques es basar la asignación de las probabilidades en el concepto simple de resultados igualmente factibles. Esto se ilustra con el ejemplo siguiente.

EJEMPLO 3-9

Se seleccionará un diodo láser, *al azar,* de un lote de 100. El espacio muestral es el conjunto de 100 diodos. La expresión "al azar" implica que es razonable suponer que cada diodo del lote tiene una posibilidad igual de ser seleccionado. Como la suma de las probabilidades debe ser igual a uno, el modelo de probabilidad para este experimento asigna la probabilidad de 0.01 a cada uno de los 100 resultados. Puede hacerse una interpretación de la probabilidad imaginando muchas repeticiones del experimento. Cada repetición empieza con los 100 diodos y se selecciona uno al azar. La probabilidad 0.01 asignada a un diodo particular representa la proporción de las repeticiones en las que se selecciona un diodo particular.

Con frecuencia, el resultado siguiente es útil.

> Siempre que un espacio muestral conste de N resultados posibles que son igualmente factibles, la probabilidad de cada resultado es $1/N$.

Muchas veces es necesario asignar probabilidades a eventos que se componen de varios resultados de un espacio muestral.

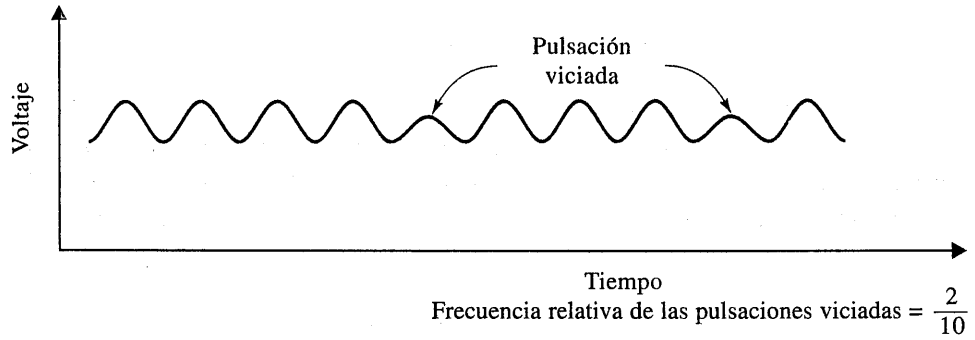


Figura 3-11 Frecuencia relativa de las pulsaciones viciadas enviadas en un canal de comunicación.

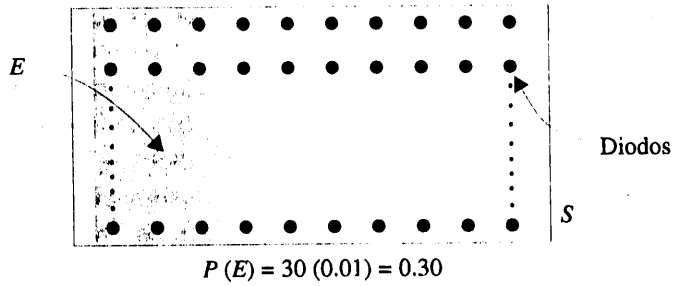


Figura 3-12 La probabilidad del evento E es la suma de las probabilidades de los resultados en E.

EJEMPLO 3-10

Suponga que 30% de los diodos láser de un lote de 100 satisface los requerimientos de energía mínimos de un cliente específico. Si se selecciona un diodo láser al azar, es decir, para cada diodo láser es igualmente posible ser seleccionado, nuestra intuición nos dice que la probabilidad de satisfacer los requerimientos del cliente es 0.30.

Sea E lo que denote el evento de que el diodo seleccionado satisface los requerimientos del cliente. Entonces E es el subconjunto de 30 diodos que satisfacen los requerimientos del cliente. Puesto que E contiene 30 resultados y cada resultado tiene una probabilidad de 0.01, se concluye que la probabilidad de E es 0.3. La conclusión coincide con nuestra intuición. En la figura 3-12 se ilustra este ejemplo.

Para un espacio muestral discreto, la probabilidad de un evento puede definirse mediante el razonamiento usado en el ejemplo anterior. La definición siguiente proporciona un método simple para obtener las probabilidades de eventos en espacios muestrales discretos.

Definición

Para un espacio muestral discreto, la *probabilidad de un evento* E, denotada como $P(E)$, es igual a la suma de las probabilidades de los resultados en E.

EJEMPLO 3-11

Un experimento aleatorio puede producir uno de los resultados $\{a, b, c, d\}$ con probabilidades 0.1, 0.3, 0.5 y 0.1, respectivamente. Sea que A denote el evento $\{a, b\}$, B el evento $\{b, c, d\}$ y C el evento $\{d\}$.

Entonces,

$$P(A) = 0.1 + 0.3 = 0.4$$
$$P(B) = 0.3 + 0.5 + 0.1 = 0.9$$
$$P(C) = 0.1$$

Asimismo, $P(A') = 0.6$, $P(B') = 0.1$ y $P(C') = 0.9$. Además, puesto que $A \cap B = \{b\}$, $P(A \cap B) = 0.3$. Puesto que $A \cup B = \{a, b, c, d\}$, $P(A \cup B) = 0.1 + 0.3 + 0.5 + 0.1 = 1$. Puesto que $A \cap C$ es el conjunto vacío, $P(A \cap C) = 0$.

EJEMPLO 3-12

En la inspección visual de un lugar dado en las obleas de un proceso de fabricación de semiconductores se obtuvo la tabla siguiente:

Numero de partículas de contaminación	**Proporción de obleas**
0	0.40
1	0.20
2	0.15
3	0.10
4	0.05
5 o más	0.10

Si se selecciona una oblea de este proceso al azar y se inspecciona el mismo lugar, ¿cuál es la probabilidad de que no contenga partículas de contaminación? La probabilidad requerida depende exclusivamente del número de partículas de contaminación presentes en el lugar dado de la oblea. Por consiguiente, es lo mismo considerar que el espacio muestral se compone de las seis categorías que resumen el número de partículas de contaminación en una oblea. Entonces, el evento de que no hay partículas en el lugar inspeccionado de la oblea, denotado como E, puede considerarse compuesto por el resultado particular $E = \{0\}$. Por lo tanto,

$$P(E) = 0.4$$

¿Cuál es la probabilidad de que una oblea contenga tres o más partículas en el lugar inspeccionado? Sea que E denote el evento de que una oblea contiene tres o más partículas en el lugar inspeccionado. Entonces, E consta de los tres resultados {3, 4, 5 o más}. Por lo tanto,

$$P(E) = 0.10 + 0.05 + 0.10 = 0.25$$

¿Cuál es la probabilidad de que una oblea contenga 0 o más de tres partículas en el lugar inspeccionado? Sea que E denote el evento de que una oblea contiene 0 o más de tres partículas en el lugar inspeccionado. Entonces, E consta de los tres resultados {0, 4, 5 o más}. Por lo tanto,

$$P(E) = 0.40 + 0.05 + 0.10 = 0.55$$

3-2.2 Axiomas de probabilidad

Ahora que se ha definido la probabilidad de un evento, los supuestos que se han hecho en cuanto a las probabilidades, pueden reunirse en un conjunto de axiomas que deben satisfacer las probabilidades en cualquier experimento aleatorio. Los axiomas aseguran que las probabilidades asignadas en un experimento pueden interpretarse como frecuencias relativas y que las asignaciones son consistentes con la comprensión intuitiva de las relaciones entre las frecuencias relativas. Por ejemplo, si el evento A está contenido en el evento B, entonces se tendrá que $P(A) \leq P(B)$. Los **axiomas no determinan las probabilidades**; las probabilidades se asignan con base en nuestro conocimiento del sistema bajo estudio. Sin embargo, los axiomas permiten calcular con facilidad las probabilidades de algunos eventos a partir del conocimiento de las probabilidades de otros eventos.

Axiomas de probabilidad

La probabilidad es un número que se asigna a cada miembro de una colección de eventos de un experimento aleatorio que satisface las siguientes propiedades.

Si S es el espacio muestral y E es cualquier evento en un experimento aleatorio,

1) $P(S) = 1$
2) $0 \leq P(E) \leq 1$
3) Para dos eventos E_1 y E_2 con $E_1 \cap E_2 = \varnothing$, $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

Los axiomas y sus consecuencias se restringen a la asignación de probabilidades en una manera que permite interpretar las probabilidades como frecuencias relativas sin inconsistencias. La propiedad de que $0 \leq P(E) \leq 1$ es equivalente al requerimiento de que una frecuencia relativa debe estar entre cero y uno. La propiedad de que $P(S) = 1$ es una consecuencia del hecho de que un resultado que proviene del espacio muestral ocurre en cada ensayo de un experimento. En consecuencia, la frecuencia relativa de S es uno. La propiedad 3 implica que si los eventos E_1 y E_2 no tienen resultados en común, entonces la frecuencia relativa de los resultados en $E_1 \cup E_2$ es la suma de las frecuencias relativas de los resultados de E_1 y E_2.

Estos axiomas implican los siguientes resultados. Las deducciones se dejan como ejercicios al final de esta sección. Ahora bien,

$$P(\varnothing) = 0$$

y para cualquier evento E,

$$P(E') = 1 - P(E)$$

Si la frecuencia relativa del evento E es 0.4, la interpretación de frecuencia relativa empleada aquí implica que la frecuencia relativa de E' es 0.6. Además, si el evento E_1 está contenido en el evento E_2, entonces

$$P(E_1) \leq P(E_2)$$

EJERCICIOS DE LA SECCIÓN 3-2

3-27. Cada uno de los cinco resultados posibles de un experimento aleatorio es igualmente factible. El espacio muestral es $\{a, b, c, d, e\}$. Sea que A denote el evento $\{a, b\}$ y sea que B denote el evento $\{c, d, e\}$. Determine las siguientes probabilidades:

a) $P(A)$
b) $P(B)$
c) $P(A')$
d) $P(A \cup B)$
e) $P(A \cap B)$

3-28. El espacio muestral de un experimento aleatorio es $\{a, b, c, d, e\}$ con probabilidades 0.1, 0.1, 0.2, 0.4 y 0.2, respectivamente. Sea que A denote el evento $\{a, b, c\}$ y sea que B denote el evento $\{c, d, e\}$. Determine las siguientes probabilidades:

a) $P(A)$
b) $P(B)$
c) $P(A')$
d) $P(A \cup B)$
e) $P(A \cap B)$

3-29. Es igualmente factible que una pieza seleccionada para prueba se haya fabricado en cualquiera de seis herramientas cortadoras.

a) ¿Cuál es el espacio muestral?
b) ¿Cuál es la probabilidad de que la pieza sea de la herramienta 1?
c) ¿Cuál es la probabilidad de que la pieza sea de la herramienta 3 o de la herramienta 5?
d) ¿Cuál es la probabilidad de que la pieza no sea de la herramienta 4?

3-30. Es igualmente factible que una pieza moldeada por inyección se obtenga de cualquiera de ocho cavidades de un molde.

a) ¿Cuál es el espacio muestral?
b) ¿Cuál es la probabilidad de que la pieza sea de la cavidad 1 o 2?
c) ¿Cuál es la probabilidad de que la pieza no sea de la cavidad 3 ni de la 4?

3-31. Un espacio muestral contiene 20 resultados igualmente factibles. Si la probabilidad del evento A es 0.3, ¿cuántos resultados hay en el evento A?

3-32. Se hace el resumen de los pedidos de un aparato de iluminación por los accesorios opcionales que se solicitan:

	proporción de pedidos
sin accesorios opcionales	0.3
con un accesorio optional	0.5
con más de un accesorio opcional	0.2

a) ¿Cuál es la probabilidad de que en un pedido se solicite al menos un accesorio opcional?
b) ¿Cuál es la probabilidad de que en un pedido no se solicite más de un accesorio opcional?

3-33. Si es igualmente factible que el último dígito de la medición de un peso sea cualquiera de los dígitos del 0 al 9,

a) ¿cuál es la probabilidad de que el último dígito sea cero?, y
b) ¿cuál es la probabilidad de que el último dígito sea mayor o igual que cinco?

3-34. El 25% de los técnicos de un laboratorio hace correctamente una preparación muestral para una medición química; 70% la hace con un error secundario y 5% la hace con un error grave.

a) Si se selecciona aleatoriamente a un técnico para hacer la preparación, ¿cuál es la probabilidad de que la haga sin error?
b) ¿Cuál es la probabilidad de que la haga con un error secundario o con uno grave?

3-35. Se analizan las placas circulares plásticas de policarbonato de un proveedor para la resistencia a las rayaduras y la resistencia a los impactos. Los resultados de 100 placas circulares se resumen a continuación.

		resistencia a los impactos	
		alta	baja
resistencia a las rayaduras	alta	80	9
	baja	6	5

Sea que A denote el evento de que una placa circular tiene alta resistencia a los impactos y sea que B denote el evento de que una pla-

ca circular tiene alta resistencia a las rayaduras. Si se selecciona una placa circular al azar, determine las siguientes probabilidades:

a) $P(A)$
b) $P(B)$
c) $P(A')$
d) $P(A \cap B)$
e) $P(A \cup B)$
f) $P(A' \cup B)$

3-36. Muestras de una pieza de aluminio forjado se clasifican con base en el acabado de la superficie (en micropulgadas) y en las mediciones de la longitud. Los resultados de 100 piezas se resumen a continuación:

		longitud	
		excelente	buena
acabado de	excelente	75	7
la superficie	bueno	10	8

Sea que A denote el evento de que una muestra tiene el terminado de la superficie excelente y sea que B denote el evento de que una muestra tiene una longitud excelente. Si se selecciona una pieza al azar, determine las siguientes probabilidades:

a) $P(A)$
b) $P(B)$
c) $P(A')$
d) $P(A \cap B)$
e) $P(A \cup B)$
f) $P(A' \cup B)$

3-37. Se clasifican muestras de hule espuma de tres proveedores de acuerdo a si cumplen o no con las especificaciones. Los resultados de 100 muestras se resumen a continuación:

		cumple	
		sí	no
	1	18	2
proveedor	2	17	3
	3	50	10

Sea que A denote el evento de que una muestra es del proveedor 1 y sea que B denote el evento de que una muestra cumple con las especificaciones. Si se selecciona una muestra de hule espuma al azar, determine las siguientes probabilidades:

a) $P(A)$
b) $P(B)$
c) $P(A')$
d) $P(A \cap B)$
e) $P(A \cup B)$
f) $P(A' \cup B)$

3-38. Aplique los axiomas de probabilidad para demostrar lo siguiente:

a) Para cualquier evento E, $P(E') = 1 - P(E)$.
b) $P(\varnothing) = 0$.
c) Si A está contenido en B, entonces $P(A) \leq P(B)$.

3-3 REGLAS DE ADICIÓN

Los eventos conjuntos se generan aplicando operaciones básicas de conjuntos a eventos individuales. Las uniones de eventos, tales como $A \cup B$, las intersecciones de eventos, tales como $A \cap B$, y los complementos de eventos, tales como A', suelen ser de interés. La probabilidad de un evento conjunto con frecuencia puede determinarse a partir de las probabilidades de los eventos individuales que lo componen. En ocasiones, las operaciones básicas con conjuntos también resultan útiles para determinar la probabilidad de un evento conjunto.

EJEMPLO 3-13

En la tabla 3-1 se enlista el historial de 940 obleas en un proceso de fabricación de semiconductores. Supóngase que se selecciona una de las obleas de la tabla 3-1 al azar. Sea que A denote el evento de que la oblea contiene niveles altos de contaminación. Entonces,

$$P(A) = 358/940$$

Tabla 3-1 Obleas en un proceso de fabricación de semiconductores

		Centro de la máquina-herramienta de deposición electrónica	
		no	sí
Contaminación	no	514	68
alta	sí	112	246

Sea que B denote el evento de que la oblea se encuentra en el centro de la máquina-herramienta de deposición electrónica. Entonces,

$$P(B) = 314/940$$

Además, $P(A \cap B)$ es la probabilidad de que la oblea sea del centro de la máquina-herramienta de deposición electrónica y contenga niveles altos de contaminación. Por la tabla 3-1,

$$P(A \cap B) = 246/940$$

El evento $A \cup B$ es el evento de que una oblea es del centro de la máquina-herramienta de deposición electrónica o contiene niveles altos de contaminación (o ambos). Por la tabla 3-1, $P(A \cup B) = 426/940$. Un cálculo alternativo de $P(A \cup B)$ puede obtenerse de la siguiente manera: las 246 obleas que componen el evento $A \cap B$ se incluyen una vez en el cálculo de $P(A)$ y una vez más en el cálculo de $P(B)$. Por lo tanto, puede encontrarse que $P(A \cup B)$ es

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 358/940 + 314/940 - 246/940 \\ &= 426/940 \end{aligned}$$

Sea E el evento de que la oblea no es del centro de la máquina-herramienta de deposición electrónica ni contiene niveles altos de contaminación. Por la tabla 3-1,

$$P(E) = 514/940$$

Como una demostración de las operaciones con conjuntos, puede emplearse un cálculo alternativo para determinar $P(E)$. Se tiene que

$$E = (A \cup B)'$$

Por lo tanto,

$$P(E) = 1 - P(A \cup B) = 514/940$$

El ejemplo anterior motiva el siguiente resultado general:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (3\text{-}1)$$

EJEMPLO 3-14

Obleas como las descritas en el ejemplo 3-13 se clasificaron adicionalmente según estuvieran en el "centro" o en el "borde" de la máquina-herramienta de deposición electrónica que se empleó en su fabricación y por el grado de contaminación. En la tabla 3-2 se muestra la proporción de obleas en cada categoría. ¿Cuál es la probabilidad de que una oblea seleccionada al azar de este lote estuviera en el centro de la máquina-herramienta de deposición electrónica? Como 0.72 de las obleas están cerca del centro, la probabilidad es 0.72.

¿Cuál es la probabilidad de que una oblea contenga cuatro o más partículas y que estuviera en el borde? Sea que E_1 denote el evento de que una oblea contiene cuatro o más partículas y sea que E_2 denote el evento de que una oblea está en el borde. La probabilidad pedida es $P(E_1 \cap E_2)$. La proporción de obleas que contienen cuatro o más partículas y que están en el borde es 0.01 + 0.03. La probabilidad pedida es 0.04.

¿Cuál es la probabilidad de que una oblea estuvo en el borde o que contenga cuatro o más partículas? Usando las definiciones de la pregunta anterior, se encuentra que la probabilidad pedida es $P(E_1 \cup E_2)$. Ahora bien, $P(E_1) = 0.15$ y $P(E_2) = 0.28$. Por la pregunta anterior, $P(E_1 \cap E_2) = 0.04$. Por lo tanto, aplicando la ecuación 3-1, se encuentra que

$$P(E_1 \cup E_2) = 0.15 + 0.28 - 0.04 = 0.39$$

¿Cuál es la probabilidad de que una oblea contenga menos de dos partículas o que sea a la vez del borde y contenga más de cuatro partículas? Sea que E_1 denote el evento de que una oblea contiene menos de dos partículas y sea que E_2 denote el evento de que una oblea es a la vez del borde y contiene más de cuatro partículas. La probabilidad pedida es $P(E_1 \cup E_2)$. Ahora bien, $P(E_1) = 0.60$ y $P(E_2) = 0.03$. Además, E_1 y E_2 son mutuamente excluyentes. Por consiguiente, no hay obleas en la intersección y $P(E_1 \cap E_2) = 0$. Por lo tanto,

$$P(E_1 \cup E_2) = 0.60 + 0.03 = 0.63$$

Tabla 3-2 Obleas clasificadas por contaminación y localización

Número de partículas de contaminación	Centro	Borde	Totales
0	0.30	0.10	0.40
1	0.15	0.05	0.20
2	0.10	0.05	0.15
3	0.06	0.04	0.10
4	0.04	0.01	0.05
5 o más	0.07	0.03	0.10
Totales	0.72	0.28	1.00

Recuérdese que se dice que dos eventos A y B son mutuamente excluyentes si $A \cap B = \varnothing$. Entonces, $P(A \cap B) = 0$, y el resultado general para la probabilidad de $A \cup B$ se reduce al tercer axioma de probabilidad.

Si A y B son eventos mutuamente excluyentes, entonces

$$P(A \cup B) = P(A) + P(B) \tag{3-2}$$

Tres o más eventos

Probabilidades más complicadas, tales como $P(A \cup B \cup C)$, pueden determinarse aplicando varias veces la ecuación 3-1 y utilizando algunas de las operaciones básicas con conjuntos. Por ejemplo,

$$P(A \cup B \cup C) = P[(A \cup B) \cup C)] = P(A \cup B) + P(C) - P[(A \cup B) \cap C]$$

Al desarrollar $P(A \cup B)$ por la ecuación 3-1 y aplicar la regla distributiva para operaciones con conjuntos a fin de simplificar $P[(A \cup B) \cap C]$, se obtiene

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) - P(A \cap B) + P(C) \\ &\quad - P[(A \cap C) \cup (B \cap C)] \\ &= P(A) + P(B) - P(A \cap B) + P(C) \\ &\quad - [P(A \cap B) + P(B \cap C) - P(A \cap B \cap C)] \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$

Se ha desarrollado una fórmula para la probabilidad de la unión de tres eventos. Es posible desarrollar una fórmula para la probabilidad de la unión de cualquier número de eventos, aunque las fórmulas se hacen bastante complejas. Como resumen, para el caso de tres eventos,

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) - P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \end{aligned} \tag{3-3}$$

En general, se dice que los eventos de una colección $E_1, E_2, \ldots, E_k$, son mutuamente excluyentes si no hay traslape entre cualquiera de ellos.

Definición

Se dice que los eventos de una colección $E_1, E_2, \ldots, E_k$ son **mutuamente excluyentes** si para todos los pares,

$$E_i \cap E_j = \varnothing$$

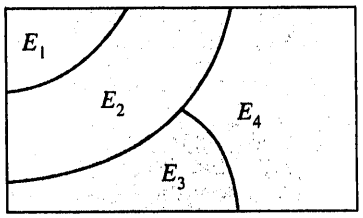


Figura 3-13 Diagrama de Venn de cuatro eventos mutuamente excluyentes.

En la figura 3-13 se muestra el diagrama de Venn de varios eventos mutuamente excluyentes. Al generalizar el razonamiento para la unión de dos eventos se puede llegar al siguiente resultado:

Para una colección de eventos mutuamente excluyentes,

$$P(E_1 \cup E_2 \cup \ldots \cup E_k) = P(E_1) + P(E_2) + \ldots P(E_k) \qquad (3\text{-}4)$$

EJEMPLO 3-15

En una operación de maquinado, sea que x denote la longitud de una pieza y supóngase que para

10% de las piezas, $x \leq 7.55$ milímetros.

15% de las piezas, $7.55 < x \leq 7.57$ milímetros.

25% de las piezas, $7.57 < x \leq 7.59$ milímetros.

Si se selecciona una pieza de esta operación, ¿cuál es la probabilidad de que sea menor o igual que 7.59 milímetros? Sea que E_1 denote el evento de que $x \leq 7.55$ milímetros de longitud. Sea que E_2 denote el evento de que $7.55 < x \leq 7.57$ milímetros. Sea que E_3 denote el evento de que $7.57 < x \leq 7.59$ milímetros. Entonces, puesto que estos eventos son mutuamente excluyentes,

$$P(E_1 \cup E_2 \cup E_3) = 0.10 + 0.15 + 0.25 = 0.50$$

EJERCICIOS DE LA SECCIÓN 3-3

3-39. Si $P(A) = 0.3$, $P(B) = 0.2$ y $P(A \cap B) = 0.1$, determine las siguientes probabilidades:

a) $P(A')$
b) $P(A \cup B)$
c) $P(A' \cap B)$
d) $P(A \cap B')$
e) $P[(A \cup B)']$
f) $P(A' \cup B)$

3-40. Si A, B y C son eventos mutuamente excluyentes con $P(A) = 0.2$, $P(B) = 0.3$ y $P(C) = 0.4$, determine las siguientes probabilidades:

a) $P(A \cup B \cup C)$
b) $P(A \cap B \cap C)$
c) $P(A \cap B)$
d) $P[(A \cup B) \cap C]$

3-41. Si A, B y C son eventos mutuamente excluyentes, ¿es posible que $P(A) = 0.3$, $P(B) = 0.4$ y $P(C) = 0.5$? ¿Por qué sí o por qué no?

3-42. Se analizan las placas circulares plásticas de policarbonato de un proveedor para la resistencia a las rayaduras y la resistencia a los impactos. Los resultados de 100 placas circulares se resumen a continuación:

		resistencia a los impactos	
		alta	baja
resistencia a	alta	80	9
las rayaduras	baja	6	5

a) Si se selecciona una placa circular al azar, ¿cuál es la probabilidad de que su resistencia a las rayaduras sea alta y su resistencia a los impactos sea alta?

b) Si se selecciona una placa circular al azar, ¿cuál es la probabilidad de que su resistencia a las rayaduras sea alta o su resistencia a los impactos sea alta?

c) Considérese el evento de que una placa circular tiene alta resistencia a las rayaduras y el evento de que una placa circular tiene alta resistencia a los impactos. ¿Estos dos eventos son mutuamente excluyentes?

3-43. El análisis de las flechas para un compresor se resumen por su cumplimiento con las especificaciones:

		la redondez cumple	
		sí	no
el acabado de la	sí	345	5
superficie cumple	no	12	8

a) Si se selecciona una flecha al azar, ¿cuál es la probabilidad de que la flecha cumpla con los requerimientos del acabado de la superficie?

b) ¿Cuál es la probabilidad de que la flecha seleccionada cumpla con los requerimientos del acabado de la superficie o con los requerimientos de redondez?

c) ¿Cuál es la probabilidad de que la flecha seleccionada cumpla con los requerimientos del acabado de la superficie o bien de que no cumpla con los requerimientos de redondez?

d) ¿Cuál es la probabilidad de que la flecha seleccionada cumpla tanto con los requerimientos del acabado de la superficie como con los requerimientos de redondez?

3-44. Las flechas del ejercicio 3-43 se clasifican adicionalmente en términos de la máquina-herramienta que se usó para fabricarlas.

Herramienta 1

		la redondez cumple	
		sí	no
el acabado de la	sí	200	1
superficie cumple	no	4	2

Herramienta 2

		la redondez cumple	
		sí	no
el acabado de la	sí	145	4
superficie cumple	no	8	6

a) Si se selecciona una flecha al azar, ¿cuál es la probabilidad de que la flecha cumpla con los requerimientos del acabado de la superficie o con los requerimientos de redondez, o sea, de la máquina-herramienta 1?

b) Si se selecciona una flecha al azar, ¿cuál es la probabilidad de que la flecha cumpla con los requerimientos del acabado de la superficie o no cumpla con los requerimientos de redondez, o sea, de la máquina-herramienta 2?

c) Si se selecciona una flecha al azar, ¿cuál es la probabilidad de que la flecha cumpla tanto con los requerimientos del acabado de la superficie como con los de redondez, o que la flecha sea de la máquina-herramienta 2?

d) Si se selecciona una flecha al azar, ¿cuál es la probabilidad de que la flecha cumpla con los requerimientos del acabado de la superficie, o que la flecha sea de la máquina-herramienta 2?

3-4 PROBABILIDAD CONDICIONAL

Los siguientes ejemplos ilustran la necesidad de reevaluar las probabilidades cuando se dispone de información adicional.

Un canal de comunicación digital tiene un índice de error de un bit por cada mil bits transmitidos. Los errores son raros, pero cuando ocurren tienden a hacerlo en rachas que afectan varios bits consecutivos. Si se transmite un solo bit, un modelo de la probabilidad de un error sería 1/1 000. Sin embargo, si el bit previo tuvo un error, debido a una racha, podría conjeturarse que la probabilidad de que el siguiente bit tenga un error es mayor que 1/1 000.

En el proceso de fabricación de una película delgada, la proporción de partes que no son aceptables es 2%. Sin embargo, el proceso es sensible a problemas de contaminación que pueden aumentar el porcentaje de partes que no son aceptables. Si se sabe que durante un turno dado hubo problemas con los filtros empleados para controlar la contaminación, la probabilidad de que una parte sea inaceptable se tasaría en un valor mayor que 2%.

EJEMPLO 3-16

En un proceso de fabricación, 10% de las piezas presentan imperfecciones superficiales visibles y 25% de las piezas con imperfecciones superficiales son funcionalmente defectuosas. Sin embargo, sólo 5% de las piezas sin imperfecciones superficiales son funcionalmente defectuosas. La probabilidad de una pieza con un defecto funcional depende del conocimiento de la presencia o ausencia de una imperfección superficial. Si una pieza tiene una imperfección superficial, la probabilidad de que sea funcionalmente defectuosa es 0.25. Si una pieza no tiene ninguna imperfección superficial, la probabilidad de que sea funcionalmente defectuosa es 0.05.

Continuando con la discusión del ejemplo anterior, sea que D denote el evento de que una pieza es funcionalmente defectuosa y sea que F denote el evento de que una pieza tiene una imperfección superficial. Entonces, la probabilidad de D dado que, o suponiendo que una pieza tiene una imperfección superficial, se denota como $P(D \mid F)$. Esta notación se lee como la **probabilidad condicional** de D dado F y se interpreta como la probabilidad de que una pieza sea funcionalmente defectuosa, dado que la pieza tiene una imperfección superficial. Como 25% de las piezas con imperfecciones superficiales son funcionalmente defectuosas, la conclusión puede enunciarse como $P(D \mid F) = 0.25$. Además, puesto que F' denota el evento de que una pieza no tiene una imperfección superficial, y ya que 5% de las piezas sin imperfecciones superficiales son funcionalmente defectuosas, se tiene que $P(D \mid F') = 0.05$. Estos resultados se muestran gráficamente en la figura 3-14.

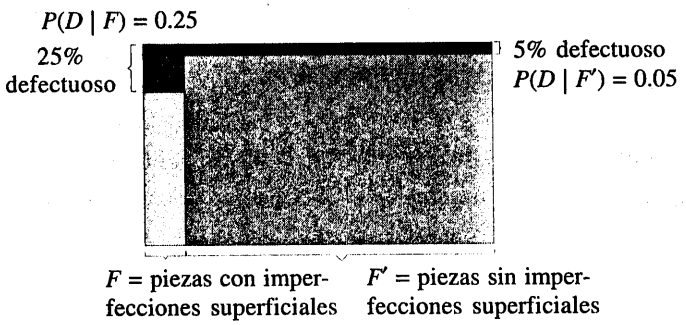


Figura 3-14 Probabilidades condicionales para las piezas con imperfecciones superficiales.

En algunos modelos, $P(B \mid A)$ pueden calcularse directamente de la descripción del experimento aleatorio.

EJEMPLO 3-17

La producción de un día de 850 partes manufacturadas contiene 50 partes que no satisfacen los requerimientos del cliente. Se seleccionan dos partes *al azar,* sin reemplazos, del lote. La expresión "al azar" implica que, cuando se selecciona la primera parte, todas las partes son igualmente factibles, y cuando se selecciona la segunda, todas las partes restantes son igualmente factibles.

¿Cuál es la probabilidad de que la segunda parte sea defectuosa dado que la primera es defectuosa? Sea que A denote el evento de que la primera parte seleccionada es defectuosa y sea que B denote el evento de que la segunda parte seleccionada es defectuosa. La probabilidad que se pide puede expresarse como $P(B \mid A)$.

Si la primera parte es defectuosa, entonces antes de seleccionar la segunda parte el lote contiene 849 partes, de las cuales 49 son defectuosas. Por lo tanto,

$$P(B \mid A) = 49/849$$

EJEMPLO 3-18

Continuando con el ejemplo anterior, si se seleccionan tres partes al azar, ¿cuál es la probabilidad de que las dos primeras sean defectuosas y la tercera no lo sea? Este evento puede describirse en notación abreviada, simplemente como $P(ddn)$. Se tiene

$$P(ddn) = \frac{50}{850} \cdot \frac{49}{849} \cdot \frac{800}{848} = 0.0032$$

El tercer término se obtiene de la siguiente manera: después de seleccionar las dos primeras partes, quedan 848; de las partes que quedan, 800 no son defectuosas. En este ejemplo, es fácil obtener la solución interpretando la probabilidad condicional.

Como lo ilustran los ejemplos anteriores, con frecuencia la estimación de la probabilidad de un evento se actualiza como resultado de la información adicional. Las probabilidades condicionales pueden encontrarse a partir de la definición formal de probabilidad condicional.

Definición

La **probabilidad condicional** de un evento B dado un evento A, denotada como $P(B \mid A)$, es

$$P(B \mid A) = P(A \cap B) / P(A) \tag{3-5}$$

para $P(A) > 0$.

Esta definición puede entenderse en el caso especial en que todos los resultados de un experimento aleatorio son igualmente factibles. Si hay n resultados totales, entonces

$$P(A) = (\text{número de resultados en } A)/n$$

Además,

$$P(A \cap B) = (\text{número de resultados en } A \cap B)/n$$

Por consiguiente,

$$P(A \cap B)/P(A) = (\text{número de resultados en } A \cap B)/(\text{número de resultados en } A)$$

Por lo tanto, $P(B \mid A)$ puede interpretarse como la frecuencia relativa del evento B entre los ensayos que producen un resultado en el evento A.

EJEMPLO 3-19

El registro de 266 muestras de aire se ha clasificado con base en la presencia de dos moléculas raras. Sea que A denote el evento que consta de todas las muestras de aire en las que está presente la molécula rara 1 y sea que B denote el evento que consta de todas las muestras de aire en las que está presente la molécula 2. Utilizando los resultados de la tabla 3-3, se encuentra que

$$\begin{aligned} &\text{P(molécula 2 presente} \mid \text{molécula 1 presente)} \\ &= P(B \mid A) \\ &= P(A \cap B)/P(A) \\ &= \frac{12}{266} \div \frac{36}{266} = 12/36 \end{aligned}$$

Obsérvese que en este ejemplo las cuatro probabilidades siguientes son diferentes:

$$P(A) = 36/266, \qquad P(A \mid B) = 12/30$$

$$P(B) = 30/266, \qquad P(A \mid B) = 12/36$$

Aquí, $P(A)$ y $P(A \mid B)$ son las probabilidades del mismo evento, pero se calculan bajo dos condiciones de conocimiento diferentes. De manera similar, $P(B)$ y $P(B \mid A)$ son las probabilidades del mismo evento, pero calculadas bajo dos condiciones de conocimiento diferentes.

Resultados similares se muestran en el diagrama de árbol de la figura 3-15.

Tabla 3-3 Moléculas en muestras de aire

		Molécula 1 presente		
		no	sí	
Molécula 2 presente	no	212	24	236
	sí	18	12	30
		230	36	266

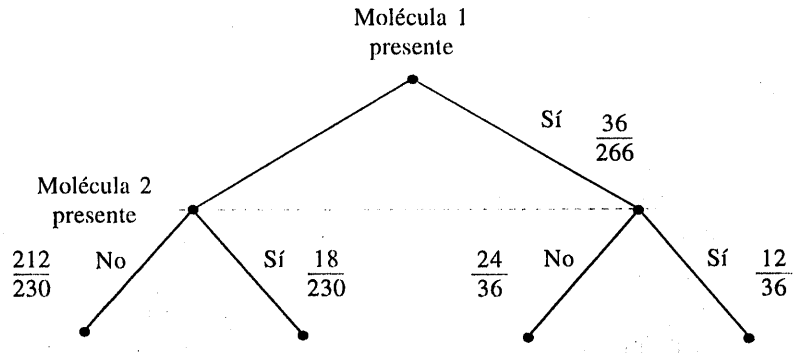


Figura 3-15 Diagrama de árbol para las moléculas en las muestras de aire.

EJERCICIOS DE LA SECCIÓN 3-4

3-45. Se analizan las placas circulares plásticas de policarbonato de un proveedor para la resistencia a las rayaduras y la resistencia a los impactos. Los resultados de 100 placas circulares se resumen a continuación.

		resistencia a los impactos	
		alta	baja
resistencia a	alta	80	9
las rayaduras	baja	6	5

Sea que A denote el evento de que una placa circular tiene alta resistencia a los impactos y sea que B denote el evento de que una placa circular tiene alta resistencia a las rayaduras. Determine las siguientes probabilidades:

a) $P(A)$
b) $P(B)$
c) $P(A \mid B)$
d) $P(B \mid A)$

3-46. Se clasifican muestras de aluminio fundido con base en el acabado de la superficie (en micropulgadas) y las mediciones de la longitud. Los resultados de 100 piezas se resumen a continuación:

		longitud	
		excelente	buena
acabado de	excelente	75	7
la superficie	bueno	10	8

Sea que A denote el evento de que una muestra tiene acabado de la superficie excelente y sea que B denote el evento de que una muestra tenga una longitud excelente. Determine:

a) $P(A)$
b) $P(B)$
c) $P(A \mid B)$
d) $P(B \mid A)$
e) Si la pieza seleccionada tiene acabado de la superficie excelente, ¿cuál es la probabilidad de que la longitud sea excelente?
f) Si la pieza seleccionada tiene buena longitud, ¿cuál es la probabilidad de que el acabado de la superficie sea excelente?

3-47. El análisis de las flechas para un compresor se resumen por su cumplimiento con las especificaciones:

		la redondez cumple	
		sí	no
el acabado de la	sí	345	5
superficie cumple	no	12	8

a) Si se sabe que una flecha cumple con los requerimientos de redondez, ¿cuál es la probabilidad de que cumpla con los requerimientos del acabado de la superficie?
b) Si se sabe que una flecha no cumple con los requerimientos de redondez, ¿cuál es la probabilidad de que cumpla con los requerimientos del acabado de la superficie?

3-48. En la tabla siguiente se resume el análisis de muestras de acero galvanizado, por el peso del recubrimiento y la rugosidad de la superficie.

		peso del recubrimiento	
		alto	bajo
rugosidad de la superficie	alta	12	16
	baja	88	34

a) Si el peso del recubrimiento de una muestra es alto, ¿cuál es la probabilidad de que la rugosidad de la superficie sea alta?
b) Si la rugosidad de la superficie de una muestra es alta, ¿cuál es la probabilidad de que el peso del recubrimiento sea alto?
c) Si la rugosidad de la superficie de una muestra es baja, ¿cuál es la probabilidad de que el peso del recubrimiento sea bajo?

3-49. Considere los datos sobre la contaminación de las obleas y la localización en la máquina-herramienta de deposición electrónica presentados en el ejemplo 3-14. Suponga que se selecciona una oblea al azar de este conjunto. Sea que A denote el evento de que una oblea contiene cuatro o más partículas y sea que B denote el evento de que es del centro de la máquina-herramienta de deposición electrónica. Determine:
a) $P(A)$
b) $P(A \mid B)$
c) $P(B)$
d) $P(B \mid A)$
e) $P(A \cap B)$
f) $P(A \cup B)$

3-50. Un lote de 100 chips semiconductores contiene 20 que están defectuosos. Se seleccionan dos chips del lote, al azar, sin reemplazo.
a) ¿Cuál es la probabilidad de que el primero que se seleccione esté defectuoso?
b) ¿Cuál es la probabilidad de que el segundo que se seleccione esté defectuoso, dado que el primero estuvo defectuoso?
c) ¿Cómo cambia la respuesta del inciso b) si los chips seleccionados se reemplazaron antes de la siguiente selección?

3-51. Un lote contiene 15 piezas fundidas de un proveedor local y 25 piezas fundidas de un proveedor del estado contiguo. Se seleccionandos piezas fundidas al azar, sin reemplazo, del lote de 40. Sea que A denote el evento de que la primera pieza fundida seleccionada es del proveedor local y sea que B denote el evento de que la segunda pieza fundida seleccionada es del proveedor local. Determine:
a) $P(A)$
b) $P(B \mid A)$
c) $P(A \cap B)$
d) $P(A \cup B)$

Continuación del ejercicio 3-51. Suponga que se seleccionan tres piezas fundidas al azar, sin reemplazo, del lote de 40. Además de las definiciones de los eventos A y B, sea que C denote el evento de que la tercera pieza fundida seleccionada es del proveedor local. Determine:
e) $P(A \cap B \cap C)$
f) $P(A \cap B \cap C')$

3-52. Un lote de 500 contenedores de jugo de naranja congelado contiene 5 que están defectuosos. Del lote se seleccionan dos, al azar y sin reemplazo. Determine:
a) ¿Cuál es la probabilidad de que el segundo contenedor seleccionado esté defectuoso, dado que el primero estuvo defectuoso?
b) ¿Cuál es la probabilidad de que ambos estén defectuosos?
c) ¿Cuál es la probabilidad que ambos sean aceptables?

3-53. Continuación del ejercicio 3-52. Del lote se seleccionan tres contenedores, al azar y sin reemplazo.
a) ¿Cuál es la probabilidad de que el tercer contenedor seleccionado esté defectuoso, dado que el primero y el segundo seleccionados estuvieron defectuosos?
b) ¿Cuál es la probabilidad de que el tercer contenedor seleccionado esté defectuoso, dado que el primer seleccionado estuvo defectuoso y el segundo seleccionado fue aceptable?
c) ¿Cuál es la probabilidad de que los tres estén defectuosos?

3-54. Si $P(A \mid B) = 1$, ¿se cumple que $A = B$? Trace un diagrama de Venn para explicar la respuesta.

3-55. Suponga que A y B son eventos mutuamente excluyentes. Construya un diagrama de Venn que contenga los tres eventos A, B y C, tales que $P(A \mid C) = 1$ y $P(B \mid C) = 0$.

3-5 REGLAS DE MULTIPLICACIÓN Y DE PROBABILIDAD TOTAL

3-5.1 Regla de multiplicación

La definición de probabilidad condicional de la ecuación 3-5 puede reescribirse para proporcionar una expresión general de la probabilidad de la intersección de dos eventos.

Regla de multiplicación

$$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A) \tag{3-6}$$

Las expresiones de la ecuación 3-6 se obtienen intercambiando A y B.

EJEMPLO 3-20

La probabilidad de que la batería de un automóvil sometida a alta temperatura en el compartimiento del motor tenga una corriente de carga baja es 0.7. La probabilidad de que la batería esté sometida a alta temperatura en el compartimiento del motor es 0.05.

Sea que A denote el evento de que una batería tiene una corriente de carga baja y sea que B denote el evento de que una batería está sometida a alta temperatura en el compartimiento del motor. La probabilidad de que una batería tenga una corriente de carga baja y esté sometida a alta temperatura en el compartimiento del motor es

$$P(A \cap B) = P(A \mid B)P(B) = 0.7 \times 0.05 = 0.035$$

3-5.2 Regla de probabilidad total

La regla de multiplicación es útil para determinar la probabilidad de un evento que depende de otros eventos. Por ejemplo, supóngase que en la fabricación de semiconductores, la probabilidad de que un chip que está sujeto a niveles de contaminación altos durante la fabricación, ocasione la falla de un producto es de 0.10. La probabilidad de que un chip que no está sujeto a niveles de contaminación altos durante la fabricación ocasione una falla en el producto es 0.005. En una corrida particular de producción, 20% de los chips están sujetos a niveles de contaminación altos. ¿Cuál es la probabilidad de que falle un producto que utilice uno de estos chips?

Evidentemente, la probabilidad requerida depende de si el chip estuvo expuesto o no a niveles de contaminación altos. Este problema puede resolverse mediante el razonamiento siguiente.

Para un evento B cualquiera, B puede expresarse como la unión de la parte de B que está en A y la parte de B que está en A'. Es decir,

$$B = (A \cap B) \cup (A' \cap B)$$

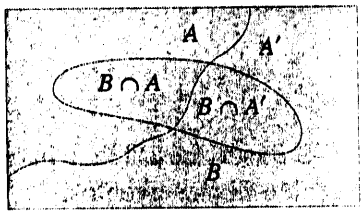


Figura 3-16 Partición de un evento en dos subconjuntos mutuamente excluyentes.

Este resultado se ilustra en el diagrama de Venn de la figura 3-16. Puesto que A y A' son mutuamente excluyentes, $A \cap B$ y $A' \cap B$ son mutuamente excluyentes. Por lo tanto, al utilizar el resultado para la probabilidad de la unión de eventos mutuamente excluyentes de la ecuación 3-2 y la regla de multiplicación de la ecuación 3-6, se obtiene la siguiente regla:

Regla de probabilidad total (para dos eventos)

Para cualesquiera eventos A y B,

$$P(B) = P(B \cap A) + P(B \cap A') = P(B \mid A)P(A) + P(B \mid A')P(A') \qquad (3\text{-}7)$$

EJEMPLO 3-21

Considérese el tema de la contaminación del principio de esta sección. Sea que F denote el evento de que el producto falla y sea que A denote el evento de que el chip está expuesto a un nivel de contaminación alto. La probabilidad requerida es $P(F)$ y la información proporcionada puede representarse como

$$P(F \mid A) = 0.10$$

$$P(F \mid A') = 0.005$$

$$P(A) = 0.20$$

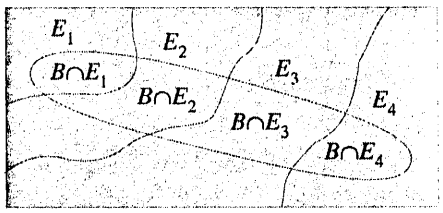
Por lo tanto,

$$P(A') = 0.80$$

al utilizar la ecuación 3-7, se obtiene

$$P(F) = 0.10(0.20) + 0.005(0.80) = 0.024$$

que puede interpretarse como el promedio ponderado de las dos probabilidades de falla.



$B = (B \cap E_1) \cup (B \cap E_2) \cup (B \cap E_3) \cup (B \cap E_4)$

$P(B) = P(B \cap E_1) + P(B \cap E_2) + P(B \cap E_3) + P(B \cap E_4)$

Figura 3-17 Partición de un evento en varios subconjuntos mutuamente excluyentes.

El razonamiento empleado para desarrollar la ecuación 3-7 puede aplicarse de manera más general. En el desarrollo de la ecuación 3-7 sólo se usaron los dos eventos mutuamente excluyentes A y A'. Sin embargo, el hecho de que $A \cup A' = S$, el espacio muestral completo, fue importante. En general, se dice que una colección de conjuntos $E_1, E_2, ..., E_k$ tales que $E_1 \cup E_2 \cup ... \cup E_k = S$ es **exhaustiva**. En la figura 3-17 se presenta una ilustración gráfica de la partición de un evento B entre una colección de eventos mutuamente excluyentes y exhaustivos.

Regla de probabilidad total (para eventos múltiples)

Suponga que $E_1, E_2, ..., E_k$ son k conjuntos mutuamente excluyentes y exhaustivos. Entonces,

$$P(B) = P(B \cap E_1) + P(B \cap E_2) + \ldots + P(B \cap E_k)$$

$$= P(B \mid E_1)\, P(E_1) + P(B \mid E_2)\, P(E_2) + \ldots + P(B \mid E_k)\, P(E_k) \qquad (3\text{-}8)$$

EJEMPLO 3-22

Continuando con el ejemplo de la fabricación de semiconductores, suponga que la probabilidad de que un chip sujeto a niveles de contaminación altos durante la fabricación, ocasione la falla de un producto, es 0.10;

la probabilidad de que un chip sujeto a niveles de contaminación medios durante la fabricación, ocasione la falla de un producto, es 0.01, y

la probabilidad de que un chip sujeto a niveles de contaminación bajos durante la fabricación ocasione la falla de un producto es 0.001.

En una corrida particular de producción, 20% de los chips están sujetos a niveles de contaminación altos, 30% a niveles de contaminación medios y 50% a niveles de contaminación bajos. ¿Cuál es la probabilidad de que un producto que use uno de estos chips falle? Sea que

H denote el evento de que un chip está expuesto a niveles de contaminación altos

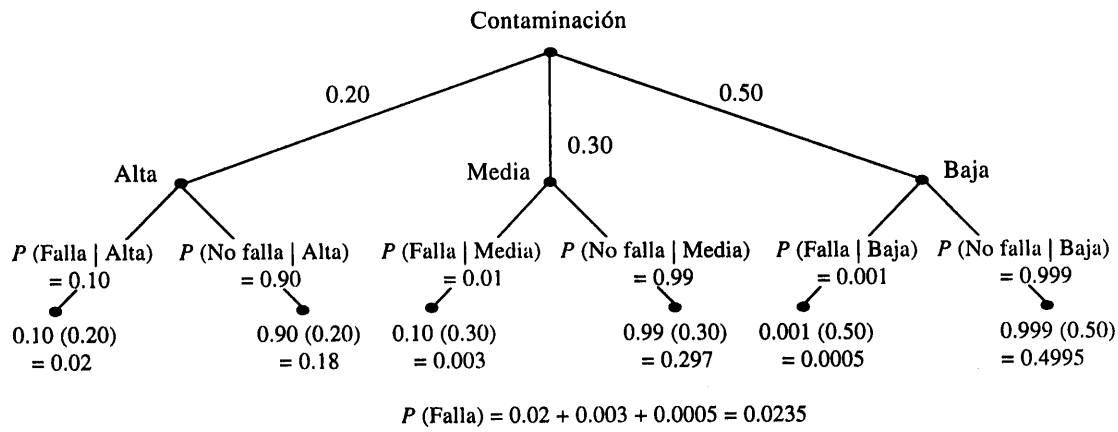


Figura 3-18 Diagrama de árbol para el ejemplo 3-22.

M denote el evento de que un chip está expuesto a niveles de contaminación medios

L denote el evento de que un chip está expuesto a niveles de contaminación bajos

Entonces,

$$\begin{aligned} P(F) &= P(F \mid H)P(H) + P(F \mid M)P(M) + P(F \mid L)P(L) \\ &= 0.10(0.20) + 0.01(0.30) + 0.001(0.50) \\ &= 0.0235 \end{aligned}$$

Este problema también se resuelve convenientemente utilizando el diagrama de árbol de la figura 3-18.

EJERCICIOS DE LA SECCIÓN 3-5

3-56. Suponga que $P(A \mid B) = 0.4$ y que $P(B) = 0.5$. Determine las siguientes probabilidades:
a) $P(A \cap B)$
b) $P(A' \cap B)$

3-57. Suponga que $P(A \mid B) = 0.2$, que $P(A \mid B') = 0.3$ y que $P(B) = 0.8$. ¿Cuál es $P(A)$?

3-58. La probabilidad de que un conector eléctrico que se mantenga seco falle durante el periodo de garantía de una computadora portátil es 1%. Si el conector siempre está húmedo, la probabilidad de una falla durante el periodo de garantía es 5%. Si 90% de los conectores se mantienen secos y 10% están húmedos, ¿qué proporción de conectores fallará durante el periodo de garantía?

3-59. Suponga que 2% de los rollos de tela de algodón y 3% de los rollos de tela de nylon contienen defectos. De los rollos usados por un fabricante, 70% son de algodón y 30% son de nylon. ¿Cuál es la probabilidad de que uno de los rollos usados por el fabricante seleccionado al azar contenga defectos?

3-60. En la fabricación de un adhesivo químico, 3% de todas las corridas de producción tienen materias primas de dos lotes diferentes. Esto ocurre cuando se rellenan los tanques de almacenamiento y la porción restante de un lote es insuficiente para llenar los tanques.

Sólo 5% de las corridas con materias primas de un solo lote requieren reprocesamiento. Sin embargo, es más difícil controlar la viscosidad de las corridas en que se usan dos o más lotes de materia prima, y 40% de las mismas requieren procesamiento adicio-

nal a fin de conseguir la viscosidad requerida.

Sea que A denote el evento de que una carga de producción está constituida por dos lotes diferentes y sea que B denote el evento de que un lote requiere procesamiento adicional. Determine las siguientes probabilidades:

a) $P(A)$
b) $P(A')$
c) $P(B \mid A)$
d) $P(B \mid A')$
e) $P(A \cap B)$
f) $P(A \cap B')$
g) $P(B)$

3-61. La rugosidad en los bordes de los productos de papel cortado aumenta con el desgaste de las cuchillas. Sólo 1% de los productos cortados con cuchillas nuevas tiene bordes rugosos, 3% de los productos cortados con cuchillas con filo promedio presentan rugosidad y 5% de los productos cortados con cuchillas desgastadas presentan rugosidad. Si 25% de las cuchillas utilizadas son nuevas, 60% tienen filo promedio y 15% están desgastadas, ¿cuál es la proporción de productos que presenta rugosidad en los bordes?

3-62. Muestras de cristal para laboratorio se colocan en paquetes pequeños y ligeros o en paquetes grandes y pesados. Suponga que 2% y 1% de las muestras enviadas en paquetes pequeños y grandes, respectivamente, se rompen durante la transportación. Si 60% de las muestras se envían en paquetes grandes y 40% se envían en paquetes pequeños, ¿qué proporción de las muestras se rompe durante la transportación?

3-63. Un lote de 25 piezas moldeadas por inyección contiene cinco que presentan una contracción excesiva.

a) Si se seleccionan dos piezas al azar, y sin reemplazo, ¿cuál es la probabilidad de que la segunda pieza seleccionada sea una con contracción excesiva?
b) Si se seleccionan tres piezas al azar, y sin reemplazo, ¿cuál es la probabilidad de que la tercera pieza seleccionada sea una con contracción excesiva?

3-64. Un lote de 100 chips semiconductores contiene 20 que están defectuosos.

a) Se seleccionan dos chips, al azar y sin reemplazo, del lote. Use la regla de probabilidad total para determinar la probabilidad de que el segundo chip seleccionado esté defectuoso.
b) Se seleccionan tres chips, al azar, y sin reemplazo, del lote. Use la regla de probabilidad total para determinar la probabilidad de que el tercer chip seleccionado esté defectuoso.

3-6 INDEPENDENCIA

En algunos casos, la probabilidad condicional $P(B \mid A)$ podría ser igual a $P(B)$. En este caso especial, el hecho de saber que el resultado del experimento está en el evento A no afecta la probabilidad de que el resultado esté en el evento B.

EJEMPLO 3-23

Suponga que la producción de un día de 850 partes manufacturadas contiene 50 partes que no cumplen con los requerimientos del cliente. Suponga además que se seleccionan dos partes del lote, pero la primera parte se reemplaza antes de seleccionar la segunda.

¿Cuál es la probabilidad de que la segunda parte esté defectuosa, dado que la primera parte está defectuosa? Sea que B denote el evento de que la segunda parte seleccionada está defectuosa y sea que A denote el evento de que la primera parte seleccionada está defectuosa. La probabilidad pedida puede expresarse como $P(B \mid A)$.

Puesto que la primera parte se reemplaza, antes de seleccionar la segunda parte el lote sigue conteniendo 850 partes, de las cuales 50 están defectuosas. Por lo tanto, la probabilidad de B no depende de si la primera parte estuvo o no defectuosa. Es decir,

$$P(B \mid A) = 50/850$$

EJEMPLO 3-24

En la tabla 3-4 se describe el registro de 84 muestras de aire con base en la presencia de dos moléculas raras.

Sea que A denote el evento que consta de todas las muestras de aire en las que está presente la molécula 1 y sea que B denote el evento que consta de todas las muestras de aire en las que está presente la molécula 2. Entonces, $P(B) = 28/84 = 1/3$. Además,

$$P(B \mid A) = P(B \cap A)/P(A) = (12/84)(36/84) = 1/3$$

En este ejemplo, el hecho de saber que la molécula 1 está presente en una muestra no cambia la probabilidad de que la molécula 2 esté presente en ella. El evento B consta de la misma proporción del número total de muestras que la proporción de muestras en A.

Además, $P(A \mid B) = P(A) = 3/7$ y $P(A \cap B) = P(A)P(B) = 1/7$.

El ejemplo precedente ilustra el siguiente resultado general.

Definición

Dos eventos son **independientes** si es verdadero cualquiera de los siguientes enunciados equivalentes.

1) $P(A \mid B) = P(A)$
2) $P(B \mid A) = P(B)$
3) $P(A \cap B) = P(A)P(B)$ (3-9)

Tabla 3-4 Moléculas en muestras de aire

		Molécula 1 presente	
		no	sí
Molécula 2 presente	no	32	24
	sí	16	12

El concepto de independencia es una importante relación entre eventos y se usa ampliamente en este libro. Una relación mutuamente excluyente entre dos eventos se basa únicamente en los resultados que comprenden los eventos. Sin embargo, una relación de independencia depende del modelo de probabilidad usado para el experimento aleatorio. Con frecuencia, se supone que la independencia es parte del experimento aleatorio que describe el sistema físico bajo estudio.

En el ejemplo 3-23, la probabilidad de que la segunda parte seleccionada esté defectuosa no depende de la primera parte porque la primera parte se reemplaza. Entonces, la probabilidad del resultado (*defectuosa, defectuosa*) se calcula como el producto de las probabilidades (50/850)(50/580) = 0.0035, al aplicar la propiedad 3 de la definición de independencia.

EJEMPLO 3-25

La producción de un día de partes manufacturadas contiene 50 partes que no cumplen con los requerimientos del cliente y 800 que sí lo hacen. Se seleccionan dos partes al azar, sin reemplazo, del lote. Sea que A denote el evento de que la primera parte está defectuosa y sea que B denote el evento de que la segunda parte está defectuosa.

Se sospecha que estos dos eventos no son independientes, porque el hecho de saber que la primera parte está defectuosa sugiere que es menos factible que la segunda parte seleccionada esté defectuosa. De hecho, $P(B \mid A) = 49/849$. Ahora bien, ¿cuál es $P(B)$? Encontrar la probabilidad no condicional $P(B)$ es un tanto difícil porque es necesario considerar los valores posibles de la primera selección.

$$\begin{aligned} P(B) &= P(B \mid A)\,P(A) + P(B \mid A')P(A') \\ &= (49/849)(50/850) + (50/849)(800/850) \\ &= 50/850 \end{aligned}$$

Como resultado interesante, $P(B)$, la probabilidad no condicional de que la segunda parte seleccionada esté defectuosa, sin ninguna información de la primera parte, es igual a $P(A)$. No obstante, nuestro objetivo es determinar la independencia. Puesto que $P(B \mid A)$ no es igual a $P(B)$, los dos eventos no son independientes, como se sospechaba.

Cuando se consideran tres o más eventos, es posible ampliar la definición de independencia. Una ampliación simple que suele ser útil es el siguiente resultado general.

Definición

Los eventos $E_1, E_2, \ldots, E_n$ son independientes si y sólo si para cualquier subconjunto $E_{i1}, E_{i2}, \ldots, E_{ik}$,

$$P(E_{i1} \cap E_{i2} \cap \ldots \cap E_{ik}) = P(E_{i1}) \times P(E_{i2}) \times \ldots \times P(E_{ik}) \qquad (3\text{-}10)$$

En general, esta definición se usa para calcular la probabilidad de que ocurran varios eventos suponiendo que son independientes y cuando se conocen las probabilidades de cada evento particular. Saber si los eventos son independientes o no, por lo general se deriva de la comprensión fundamental del experimento aleatorio.

EJEMPLO 3-26

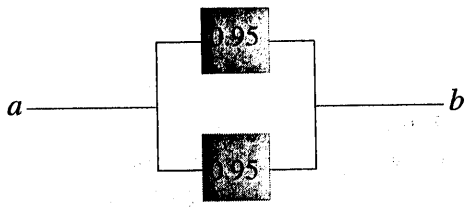
Suponga que la probabilidad de que una oblea contenga una partícula de contaminación grande es 0.01 y que las obleas son independientes; es decir, la probabilidad de que una oblea contenga una partícula grande no depende de las características de ninguna de las obleas restantes. Si se analizan 15 obleas, ¿cuál es la probabilidad de que no se encuentren partículas grandes?

Sea que E_i denote el evento de que la oblea i no contiene partículas grandes, $i = 1, 2, ..., 15$. Entonces, $P(E_i) = 0.99$. La probabilidad pedida puede representarse como $P(E_1 \cap E_2 \cap ... \cap E_{15})$. Por el supuesto de independencia y la ecuación 3-10

$$P(E_1 \cap E_2 \cap \ldots \cap E_{15}) = P(E_1) \times P(E_2) \times \ldots \times P(E_{15}) = 0.99^{15} = 0.86$$

EJEMPLO 3-27

El circuito ilustrado abajo sólo opera si hay una trayectoria de dispositivos funcionales de izquierda a derecha. La probabilidad de que cada dispositivo funcione se indica en la ilustración. Suponga que los dispositivos fallan independientemente. ¿Cuál es la probabilidad de que el circuito opere?



Sea que T y B denoten los eventos de que los dispositivos superior e inferior operan, respectivamente. Hay una trayectoria si al menos uno de los dispositivos opera. La probabilidad de que el circuito opere es

$$P(T \text{ o } B) = 1 - P[(T \text{ o } B)'] = 1 - P(T' \text{ y } B')$$

Por el supuesto de independencia,

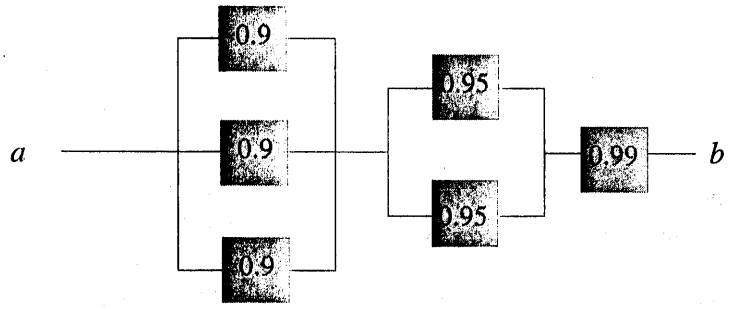
$$P(T' \text{ y } B') = P(T')P(B') = (1 - 0.95)^2 = 0.05^2$$

de donde

$$P(T \text{ o } B) = 1 - 0.05^2 = 0.9975$$

EJEMPLO 3-28

El circuito ilustrado abajo sólo opera si hay una trayectoria de dispositivos funcionales de izquierda a derecha. La probabilidad de que cada dispositivo funcione se indica en la ilustración. Suponga que los dispositivos fallan independientemente. ¿Cuál es la probabilidad de que el circuito opere?



Usando la independencia, se obtiene $(1 - 0.1^3)(1 - 0.05^2)(0.99) = 0.987$.

EJERCICIOS DE LA SECCIÓN 3-6

3-65. Si $P(A \mid B) = 0.4$, $P(B) = 0.8$ y $P(A) = 0.6$, ¿los eventos A y B son independientes?

3-66. Si $P(A \mid B) = 0.3$, $P(B) = 0.8$ y $P(A) = 0.3$, ¿los eventos B y el complemento de A son independientes?

3-67. Se analizan las placas circulares plásticas de policarbonato de un proveedor para la resistencia a las rayaduras y la resistencia a los impactos. Los resultados de 100 placas circulares se resumen a continuación:

		resistencia a los impactos	
		alta	baja
resistencia a	alta	80	9
las rayaduras	baja	6	5

Sea que A denote el evento de que una placa circular tiene alta resistencia a los impactos y sea que B denote el evento de que una placa circular tiene alta resistencia a las rayaduras. ¿Los eventos A y B son independientes?

3-68. Muestras de una pieza de aluminio fundido se clasifican con base en el acabado de la superficie (en micropulgadas) y en las mediciones de la longitud. Los resultados de 100 piezas se resumen a continuación:

		longitud	
		excelente	buena
acabado de	excelente	75	7
la superficie	bueno	10	8

Sea que A denote el evento de que una muestra tiene un terminado de la superficie excelente y sea que B denote el evento de que una muestra tiene una longitud excelente. ¿Los eventos en A y B son independientes?

3-69. Muestras de hule espuma de tres proveedores se clasifican de acuerdo a si cumplen o no con las especificaciones. Los resultados de 100 muestras se resumen a continuación:

		cumple	
		sí	no
	1	18	2
proveedor	2	17	3
	3	30	10

Sea que A denote el evento de que una muestra es del proveedor 1 y sea que B denote el evento de que una muestra cumple con las especificaciones. ¿Los eventos en A y B son independientes?

3-70. Si $P(A) = 0.2$ y $P(B) = 0.2$ y A y B son mutuamente excluyentes, ¿son independientes?

3-71. Muestras de hule espuma de dos proveedores se evalúan de acuerdo a si cumplen o no con las especificaciones. Los resultados de 126 muestras se resumen a continuación:

		cumple	
		sí	no
	1	80	4
proveedor	2	40	2

Sea que A denote el evento de que una muestra es del proveedor 1 y sea que B denote el evento de que una muestra cumple con las especificaciones.

a) ¿Son independientes A y B?

b) ¿Son independientes A' y B?

3-72. La probabilidad de que una prueba de laboratorio contenga niveles de contaminación altos es 0.10. Se verifican cinco muestras y éstas son independientes.

a) ¿Cuál es la probabilidad de que ninguna contenga niveles altos de contaminación?

b) ¿Cuál es la probabilidad de que exactamente una contenga niveles altos de contaminación?

c) ¿Cuál es la probabilidad de que al menos una contenga niveles altos de contaminación?

3-73. En una prueba de una tarjeta de circuito impreso en la que se utiliza un patrón de prueba aleatorio, un arreglo de 10 bits es igualmente factible que sea cero o uno. Suponga que los bits son independientes.

a) ¿Cuál es la probabilidad de que todos los bits sean unos?

b) ¿Cuál es la probabilidad de que todos los bits sean ceros?

c) ¿Cuál es la probabilidad de que exactamente cinco bits sean unos y cinco sean ceros?

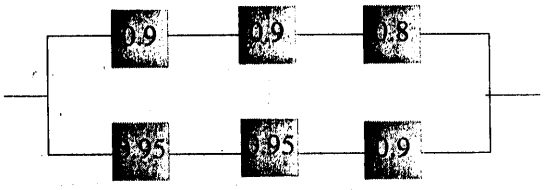
3-74. Ocho cavidades en una herramienta de moldeo por inyección produce conectores de plástico que caen en una salida común. Se escoge una muestra cada varios minutos. Suponga que las muestras son independientes.

a) ¿Cuál es la probabilidad de que cinco muestras sucesivas se hayan producido en la cavidad uno del molde?

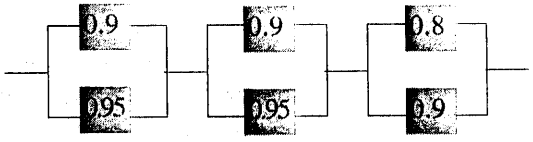
b) ¿Cuál es la probabilidad de que cinco muestras sucesivas se hayan producido en la misma cavidad del molde?

c) ¿Cuál es la probabilidad de que cuatro de cinco muestras sucesivas se hayan producido en la cavidad uno del molde?

3-75. El circuito ilustrado abajo opera si y sólo si hay una trayectoria de dispositivos funcionales de izquierda a derecha. La probabilidad de que cada dispositivo funcione se indica en la ilustración. Suponga que la probabilidad de que un dispositivo sea funcional no depende de si otros dispositivos son funcionales o no. ¿Cuál es la probabilidad de que el circuito opere?



3-76. El circuito ilustrado abajo opera si y sólo si hay una trayectoria de dispositivos funcionales de izquierda a derecha. La probabilidad de que cada dispositivo funcione se indica en la ilustración. Suponga que la probabilidad de que un dispositivo funcione no depende de si otros dispositivos son funcionales o no. ¿Cuál es la probabilidad de que el circuito opere?



3-77. Un dispositivo de almacenamiento óptico utiliza un procedimiento de recuperación de errores que requiere una lectura satisfactoria inmediata de cualquier dato escrito. Si la lectura no tiene éxito después de tres operaciones de escritura, ese sector del disco se

elimina como inaceptable para almacenamiento de datos. En una parte aceptable del disco, la probabilidad de una lectura satisfactoria es 0.98.Suponga que las lecturas son independientes. ¿Cuál es la probabilidad de que una parte aceptable del disco sea eliminada como inaceptable para almacenamiento de datos?

3-78. Un lote de 500 contenedores de jugo de naranja congelado contiene 5 que están defectuosos. Del lote se seleccionan dos, al azar y sin reemplazo. Sea que A y B denoten los eventos de que el primer y segundo contenedores están defectuosos, respectivamente.
a) ¿Son A y B eventos independientes?
b) Si el muestreo se hiciera con reemplazo, ¿serían independientes A y B?

3-7 TEOREMA DE BAYES

En algunos ejemplos, no se cuenta con una tabla completa de información como en el caso del ejemplo 3-24. Pudiera conocerse una probabilidad condicional pero nos interesaría calcular una probabilidad diferente. En el problema de la contaminación de semiconductores del ejemplo 3-22, podría formularse la siguiente pregunta. Si el chip semiconductor en el producto falla, ¿cuál es la probabilidad de que el chip se haya expuesto a niveles altos de contaminación?

Por la definición de probabilidad condicional,

$$P(A \cap B) = P(A \mid B)P(B) = P(B \cap A) = P(B \mid A)P(A)$$

Ahora bien, considerando el segundo y el último término de la expresión anterior, puede escribirse

$$P(B \mid A) = \frac{P(B \mid A)P(A)}{P(B)} \quad \text{para } P(B) > 0 \tag{3-11}$$

Este es un resultado útil que permite resolver $P(A \mid B)$ en términos de $P(B \mid A)$.

EJEMPLO 3-29

La pregunta planteada al principio de esta sección puede responderse como sigue. La probabilidad pedida puede expresarse como $P(H \mid F)$. Entonces,

$$P(H \mid F) = P(F \mid H)\, P(H) / P(F) = 0.10(0.20)/0.0235 = 0.85$$

El valor de $P(F)$ en el denominador de la solución, se encontró en el ejemplo 3-21.

En general, si $P(B)$ del denominador de la ecuación 3-11 se escribe empleando la regla de probabilidad total de la ecuación 3-8, se obtiene el siguiente resultado general que se conoce como el teorema de Bayes.

Teorema de Bayes

Si $E_1, E_2, \ldots, E_k$ son k eventos mutuamente excluyentes y exhaustivos y B es un evento cualquiera, entonces

$$P(E_1|B) = \frac{P(B|E_1)P(E_1)}{P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \ldots + P(B|E_k)P(E_k)}$$

para $P(B) > 0$

EJEMPLO 3-30

Dado que un nuevo procedimiento médico ha demostrado ser efectivo para la detección temprana de una enfermedad, se propone un estudio médico exhaustivo de la población. La probabilidad de que la prueba identifique correctamente a alguien que no padece la enfermedad como negativo es 0.95 y la probabilidad de que la prueba identifique correctamente a alguien con la enfermedad como positivo es 0.99. La incidencia de la enfermedad en la población general es 0.0001. Usted se somete a la prueba y el resultado es positivo. ¿Cuál es la probabilidad de que usted padezca la enfermedad?

Sea que D denote el evento de que usted padece la enfermedad y sea que S denote el evento de que la prueba señala positivo. La probabilidad pedida puede denotarse como $P(D \mid S)$. La probabilidad de que la prueba señale correctamente a alguien que no padezca la enfermedad como negativo es 0.95. Por consiguiente, la probabilidad de una prueba positiva en ausencia de la enfermedad es

$$P(S \mid D') = 0.05$$

Por el teorema de Bayes,

$$\begin{aligned} P(D \mid S) &= P(S \mid D)\, P(D)/[P(S \mid D)P(D) + P(S \mid D')P(D')] \\ &= 0.99(0.0001)/[0.99(0.0001) + 0.05(1 - 0.0001)] \\ &= 1/506 = 0.002 \end{aligned}$$

Sorprendentemente, aun cuando la prueba es efectiva, en el sentido de que $P(S \mid D)$ es alta y $P(S \mid D')$ es baja, debido a que la incidencia de la enfermedad en la población general es baja, las posibilidades de que usted padezca en realidad la enfermedad son muy remotas incluso si la prueba es positiva.

EJERCICIOS DE LA SECCIÓN 3-7

3-79. Suponga que $P(A \mid B) = 0.8$, $P(A) = 0.5$ y $P(B) = 0.2$. Determine $P(B \mid A)$.

3-80. El software para detectar fraudes con tarjetas telefónicas personales rastrea el número de áreas metropolitanas donde se originan las llamadas cada día. Se ha encontrado que 1% de los usuarios legítimos hacen llamadas de dos o más áreas metropolitanas en un solo

día. Sin embargo, 30% de los usuarios fraudulentos hacen llamadas de dos o más áreas metropolitanas en un solo día. La proporción de usuarios fraudulentos es 0.01%. Si el mismo usuario hace llamadas de dos o más áreas metropolitanas en un solo día, ¿cuál es la probabilidad de que el usuario sea fraudulento?

3-81. Los láser de semiconductores usados en los productos de almacenamiento óptico requieren niveles más altos de energía en las operaciones de escritura que en las de lectura. Las operaciones con niveles altos de energía reducen la vida útil del láser.

Los láser de los productos utilizados para hacer respaldos de discos magnéticos de alta velocidad realizan principalmente tareas de escritura y la probabilidad de que su vida útil rebase los cinco años es 0.95. Los láser instalados en productos que se usan para almacenamiento principal dedican aproximadamente la misma cantidad de tiempo a tareas de lectura y escritura, y la probabilidad de que su vida útil rebase los cinco años es 0.995. Ahora bien, 25% de los productos de un fabricante se usan para hacer respaldos y 75% de los productos se usan para almacenamiento principal.

Sea que A denote el evento de que la vida útil de un láser rebasa los cinco años y sea que B denote el evento de que un láser está en un producto utilizado para hacer respaldos.

Use un diagrama de árbol para determinar las siguientes probabilidades:

a) $P(B)$
b) $P(A \mid B)$
a) $P(A \mid B')$
b) $P(A \cap B)$
c) $P(A \cap B')$
d) $P(A)$
g) ¿Cuál es la probabilidad de que la vida útil de un láser rebase los cinco años?
h) ¿Cuál es la probabilidad de que un láser que falló antes de cinco años proviniera de un producto usado para hacer respaldos?

3-82. Los clientes acostumbran evaluar en forma preliminar el diseño de los productos. En el pasado, 95% de los productos de gran éxito recibieron críticas favorables, 60% de los productos con un éxito moderado recibieron críticas favorables y 10% de los productos sin mucho éxito recibieron críticas favorables. Además, 40% de los productos han sido de gran éxito, 35% han sido de éxito moderado y 25% han sido productos sin mucho éxito.

a) ¿Cuál es la probabilidad de que un producto obtenga una crítica favorable?
b) Si un diseño nuevo obtiene una crítica favorable, ¿cuál es la probabilidad de que será un producto de gran éxito?
c) Si un producto no consigue una crítica favorable, ¿cuál es la probabilidad de que será un producto de gran éxito?

3-8 VARIABLES ALEATORIAS

Con frecuencia se desea resumir el resultado de un experimento aleatorio en un solo número. En muchos de los ejemplos de experimentos aleatorios que se han considerado, el espacio muestral se ha reducido a una descripción de los resultados posibles. En algunos casos, la descripción de los resultados es suficiente, pero en otros es conveniente asociar un número con cada resultado del espacio muestral. Debido a que no se conoce de antemano el resultado particular del experimento, el valor resultante de la variable tampoco se conoce de antemano. Por esta razón, la variable que asocia un número con el resultado de un experimento aleatorio se denomina **variable aleatoria**.

Definición

Una **variable aleatoria** es una función que asigna un número real a cada resultado del espacio muestral de un experimento aleatorio.

Una variable aleatoria se denota por una letra mayúscula, como X. Se hace referencia al conjunto de los números posibles de una variable aleatoria X como el **rango de X**. Una vez que se ha llevado a cabo el experimento, el valor medido de la variable aleatoria se denota por una letra minúscula como $x = 70$ miliamperes. Con frecuencia, un experimento aleatorio se resume con el valor medido de una variable aleatoria.

Este modelo puede vincularse con los datos que se resumieron en el capítulo 2 de la siguiente manera. Los datos son los valores medidos de una variable aleatoria, los cuales se obtienen al repetir varias veces un experimento aleatorio. Por ejemplo, la primera repetición podría resultar en una medición de la corriente de $x_1 = 70.1$; al día siguiente, $x_2 = 71.2$; el tercer día, $x_3 = 71.1$; etcétera. Entonces estos datos podrán resumirse empleando los métodos descriptivos expuestos en el capítulo 2.

Con frecuencia se supone que la medición de interés —la corriente en un experimento con un alambre de cobre o la longitud de una pieza maquinada— es un número real. Entonces es posible una precisión arbitraria en las mediciones. Por supuesto, en la práctica las centésimas de una unidad podrían redondearse a la décima más próxima. Se dice que la variable aleatoria que representa esta medición es una variable aleatoria **continua**. El rango de X incluye todos los valores en un intervalo de números reales; es decir, el rango de X puede pensarse como un continuo.

En otros experimentos podría registrarse un conteo, como el del número de bits transmitidos que se reciben con error. Entonces la medición se limita a números enteros. O podría registrarse una proporción, como 0.0042 de los 10 000 bits que se recibieron fueron con error. Entonces la medición es fraccionaria, pero sigue estando limitada a puntos discretos sobre la recta real. Siempre que la medición se limita a puntos discretos sobre la recta real, se dice que la variable aleatoria es una variable aleatoria **discreta**.

Definición

Una variable aleatoria **discreta** es una variable aleatoria con un rango finito (o contablemente infinito).

Una variable aleatoria **continua** es una variable aleatoria que tiene como rango un intervalo (sea finito o infinito) de números reales.

En algunos casos, la variable aleatoria X es en realidad discreta, pero, debido a que el rango de los valores posibles es tan extenso, podría resultar más conveniente analizar X como una variable aleatoria continua. Por ejemplo, supóngase que las mediciones de la corriente se leen

en un instrumento digital que indica la corriente hasta la centésima de miliampere más próxima. Debido a que las mediciones posibles están limitadas, la variable aleatoria es discreta. Sin embargo, podría ser un enfoque más sencillo y conveniente suponer que las mediciones de la corriente son valores de una variable aleatoria continua.

Ejemplos de variables aleatorias **continuas**:
corriente eléctrica, longitud, presión, temperatura, tiempo, voltaje, peso
Ejemplos de variables aleatorias **discretas**:
número de rayaduras en una superficie, proporción de partes defectuosas entre 1 000 probadas, número de bits transmitidos recibidos con error

EJERCICIOS DE LA SECCIÓN 3-8

Decida si una variable aleatoria discreta o continua es el mejor modelo para cada una de las variables siguientes:

3-83. El tiempo hasta que un proyectil regresa a la Tierra.

3-84. El número de veces que un transistor en una memoria de computadora cambia de estado en una operación.

3-85. El volumen de gasolina que se pierde por evaporación durante el llenado de un tanque de combustible.

3-86. El diámetro exterior de una flecha maquinada.

3-87. El número de grietas que exceden media pulgada en 10 millas de una carretera interestatal.

3-88. El peso de una pieza de plástico moldeada por inyección.

3-89. El número de moléculas en una muestra de gas.

3-90. La concentración de la salida de un reactor.

3-91. La corriente en un circuito electrónico.

Ejercicios complementarios

3-92. En la prueba de los circuitos de tarjetas impresas, cada tarjeta pasa o no pasa la prueba. Una tarjeta que no pasa la prueba se somete entonces a verificación adicional para determinar cuál de los cinco tipos de defectos es el modo principal de falla. Represente el espacio muestral para este experimento.

3-93. Los datos de 200 piezas maquinadas se resumen a continuación:

	profundidad del taladro	
	mayor que la del diseño	menor que la del diseño
condición del borde		
burdo	15	10
regular	25	20
liso	60	70

a) ¿Cuál es la probabilidad de que una pieza seleccionada tenga una condición del borde regular y una profundidad del taladro menor que la del diseño?

b) ¿Cuál es la probabilidad de que una pieza seleccionada tenga una condición del borde regular o una profundidad del taladro menor que la del diseño?

c) ¿Cuál es la probabilidad de que una pieza seleccionada no tenga una condición del borde regular o no tenga una profundidad del taladro menor que la del diseño?

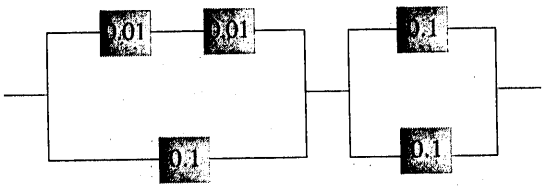
d) Construya una representación de diagrama de Venn de los eventos de este espacio muestral.

3-94. La probabilidad de que el pedido de un cliente no se envíe a tiempo es 0.05. Un cliente dado hace tres pedidos, con la suficiente separación en el tiempo como para considerarlos eventos independientes.

a) ¿Cuál es la probabilidad de que todos los pedidos se envíen a tiempo?
b) ¿Cuál es la probabilidad de que exactamente un pedido no se envíe a tiempo?
c) ¿Cuál es la probabilidad de que dos o más pedidos no se envíen a tiempo?

3-95. Una placa de acero contiene 20 pernos. Suponga que el torque de cinco pernos no está en el límite apropiado. Se seleccionan cuatro pernos al azar, sin reemplazo, para verificar el torque.
a) ¿Cuál es la probabilidad de que el torque de los cuatro pernos seleccionados esté en el límite apropiado?
b) ¿Cuál es la probabilidad de que el torque de al menos uno de los pernos seleccionados no esté en el límite apropiado?

3-96. El circuito ilustrado abajo opera si y sólo si hay una trayectoria de dispositivos funcionales de izquierda a derecha. Suponga que los dispositivos fallan independientemente y que la probabilidad de *falla* de cada dispositivo es la que se indica. ¿Cuál es la probabilidad de que el circuito opere?



3-97. En una operación de llenado automatizado, la probabilidad de un llenado incorrecto cuando el proceso opera a baja velocidad es 0.001. Cuando el proceso opera a alta velocidad, la probabilidad de un llenado incorrecto es 0.01. Suponga que 30% de los recipientes se llenan cuando el proceso opera a alta velocidad y el resto se llena cuando el proceso opera a baja velocidad.
a) ¿Cuál es la probabilidad de que un recipiente sea llenado incorrectamente?
b) Si se encuentra un recipiente llenado incorrectamente, ¿cuál es la probabilidad de que se haya llenado durante la operación a alta velocidad?

3-98. Una herramienta de encarte robotizado contiene 10 componentes principales. La probabilidad de que cualquiera de ellos falle durante el periodo de garantía es 0.01. Suponga que los componentes fallan independientemente y que la herramienta falla si cualquiera de los componentes falla. ¿Cuál es la probabilidad de que la herramienta falle durante el periodo de garantía?

3-99. Una herramienta mecánica está sin usarse 15% del tiempo. Usted solicita el uso inmediato de la herramienta mecánica en cinco ocasiones diferentes durante el año. Suponga que sus solicitudes representan eventos independientes.
a) ¿Cuál es la probabilidad de que la herramienta mecánica esté desocupada en todas las ocasiones que usted lo solicita?
b) ¿Cuál es la probabilidad de que la herramienta mecánica esté desocupada exactamente en cuatro ocasiones que usted lo solicita?
c) ¿Cuál es la probabilidad de que la herramienta mecánica esté desocupada en al menos tres ocasiones que usted lo solicita?

3-100. Un lote de 50 arandelas de separación contiene 30 arandelas cuyo grosor excede las especificaciones de diseño. Suponga que se seleccionan tres arandelas, al azar y sin reemplazo, del lote.
a) ¿Cuál es la probabilidad de que las tres arandelas seleccionadas sean más gruesas que las especificaciones del diseño?
b) ¿Cuál es la probabilidad de que la tercera arandela seleccionada sea más gruesa que las especificaciones del diseño si las dos primeras fueron más delgadas que la especificación?
c) ¿Cuál es la probabilidad de que la tercera arandela seleccionada sea más gruesa que las especificaciones del diseño?

3-101. Continuación del ejercicio 3-100. Se seleccionan arandelas del lote al azar, sin reemplazo.
a) ¿Cuál es el número mínimo de arandelas que necesitan seleccionarse para que la

probabilidad de que todas las arandelas sean más delgadas que las especificaciones de diseño sea menor que 0.10?

b) ¿Cuál es el número mínimo de arandelas que necesitan seleccionarse para que la probabilidad de que una o más arandelas sean más gruesas que las especificaciones del diseño sea al menos 0.90?

3-102. En la tabla siguiente se enlista el historial de las características opcionales en 940 pedidos de una computadora básica.

		memoria extra	
		no	sí
procesador de alta velocidad opcional	no	514	68
	sí	112	246

Sea que A denote el evento de que un pedido incluya el procesador de alta velocidad opcional y sea que B denote el evento de que un pedido requiera memoria extra. Determine las siguientes probabilidades:

a) $P(A \cup B)$

b) $P(A \cap B)$

c) $P(A' \cup B)$

d) $P(A' \cap B')$

e) ¿Cuál es la probabilidad de que un pedido requiera un procesador de alta velocidad opcional, dado que el pedido requiera memoria extra?

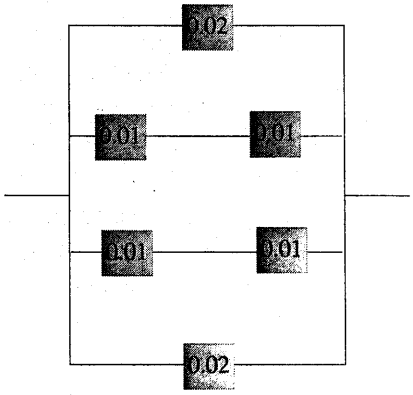
f) ¿Cuál es la probabilidad de que un pedido requiera memoria extra, dado que el pedido requiere el procesador de alta velocidad opcional?

3-103. La alineación entre la cinta magnética y la cabeza en un sistema de almacenamiento en cinta magnética, afecta el desempeño del sistema. Suponga que 10% de las operaciones de lectura sufren degradaciones por alineaciones sesgadas, 5% de las operaciones de lectura sufren degradaciones por alineaciones descentradas y las operaciones de lectura restantes tienen una alineación correcta. La probabilidad de un error de lectura es 0.01 para una alineación sesgada, 0.02 para una alineación descentrada y 0.001 para una alineación correcta.

a) ¿Cuál es la probabilidad de un error de lectura?

b) Si ocurre un error de lectura, ¿cuál es la probabilidad de que se deba a una alineación sesgada?

3-104. El circuito ilustrado abajo opera si y sólo si hay una trayectoria de dispositivos funcionales de izquierda a derecha. Suponga que los dispositivos fallan independientemente y que la probabilidad de *falla* de cada dispositivo es la que se indica. ¿Cuál es la probabilidad de que el circuito opere?



EJERCICIOS PARA DESARROLLAR EL INTELECTO

3-105. La alineación entre la cinta magnética y la cabeza en un sistema de almacenamiento de cinta magnética afecta el desempeño del sistema. Suponga que 10% de las operaciones de lectura sufren degradaciones por alineaciones sesgadas, 5% por alineaciones descentradas, 1% por alineación tanto sesgada como descentrada y las operaciones de lectura restantes tienen la alineación correcta. La probabilidad de un error de lectura es 0.01 para una alineación sesgada, 0.02 para una alineación descentrada, 0.06 para ambas condiciones y 0.001 para una alineación correcta. ¿Cuál es la probabilidad de un error de lectura?

3-106. Suponga que un lote de arandelas es lo suficientemente grande para que pueda suponerse que el muestreo se hace con reemplazo. Suponga que 60% de las arandelas exceden el grosor de diseño.

a) ¿Cuál es el número mínimo de arandelas que es necesario seleccionar para que la probabilidad de que todas las arandelas sean más delgadas que el grosor de diseño sea menor que 0.10?

b) ¿Cuál es el número mínimo de arandelas que es necesario seleccionar para que la probabilidad de que una o más arandelas sean más gruesas que el grosor de diseño sea al menos 0.90?

3-107. Una fábrica de productos biotecnológicos puede producir paquetes para pruebas de diagnóstico a un costo de $20. Cada paquete para el que hay demanda en la semana de producción puede venderse a $100. Sin embargo, la vida media de los componentes del paquete requiere que éste se deseche si no se vende en la semana de producción. El costo de desechar el paquete es de $5. La demanda semanal se resume a continuación:

Demanda semanal				
Número de unidades	0	50	100	200
Probabilidad de la demanda	0.05	0.4	0.3	0.25

¿Cuántos paquetes deberán producirse cada semana para maximizar las ganancias promedio de la fábrica?

3-108. Suponer las siguientes características del proceso de inspección del ejercicio 3-95. Si un operador verifica un perno, la probabilidad de que se identifique un perno con el torque incorrecto es 0.95. Si un perno inspeccionado tiene el torque correcto, la conclusión del operador siempre es acertada. ¿Cuál es la probabilidad de que al menos un perno de la muestra de cuatro se identifique como un perno con el torque incorrecto?

3-109. Si los eventos A y B son independientes, demuestre que A' y B' son independientes.

3-110. Suponga que la tabla del ejercicio 3-71 se generaliza de la siguiente manera:

		cumple	
		sí	no
	1	ka	kb
proveedor	2	a	b

donde a, b y k son enteros positivos. Sea que A denote el evento de que una muestra de hule espuma es del proveedor 1 y sea que B denote el evento de que una muestra de hule espuma cumple con las especificaciones. Demuestre que A y B son eventos independientes.

Este ejercicio ilustra el resultado de que siempre que los renglones de una tabla (con r renglones y c columnas) son proporcionales, un evento definido por una categoría de renglón y un evento definido por una categoría de columna son independientes.

4 Variables aleatorias discretas y distribuciones de probabilidad

PLAN GENERAL DEL CAPÍTULO

4-1 VARIABLES ALEATORIAS DISCRETAS

4-2 DISTRIBUCIONES DE PROBABILIDAD Y FUNCIONES DE MASA DE PROBABILIDAD

4-3 FUNCIONES DE DISTRIBUCIÓN ACUMULADA

4-4 MEDIA Y VARIANZA DE UNA VARIABLE ALEATORIA DISCRETA

4-5 DISTRIBUCIÓN DISCRETA UNIFORME

4-6 DISTRIBUCIÓN BINOMIAL

4-7 DISTRIBUCIONES GEOMÉTRICA Y BINOMIAL NEGATIVA

4-7.1 Distribución geométrica

4-7.2 Distribución binomial negativa

4-8 DISTRIBUCIÓN HIPERGEOMÉTRICA

4-9 DISTRIBUCIÓN DE POISSON

4-1 VARIABLES ALEATORIAS DISCRETAS

En este capítulo se presentan los modelos de probabilidad comunes para variables aleatorias discretas.

EJEMPLO 4-1

Un sistema de comunicación de voz de un negocio contiene 48 líneas externas. En un tiempo particular, se observa el sistema y algunas de las líneas están en uso. Sea que la variable aleatoria X denote el número de líneas en uso. Entonces, X puede asumir cualquiera de los valores enteros 0 al 48. Cuando el sistema se observa, si 10 líneas están en uso entonces $x = 10$.

Tabla 4-1 Pruebas de obleas

Resultado	Probabilidad	x
(*pasa, pasa*)	0.64	2
(*falla, pasa*)	0.16	1
(*pasa, falla*)	0.16	1
(*falla, falla*)	0.04	0

EJEMPLO 4-2

En un proceso de fabricación de semiconductores, se prueban dos obleas de un lote. Cada oblea se clasifica como *pasa* o *falla*. Suponga que la probabilidad de que una oblea pase la prueba es 0.8 y que las obleas son independientes. El espacio muestral del experimento y las probabilidades asociadas se muestran en la tabla 4-1. Por ejemplo, debido a la independencia, la probabilidad del resultado de que la primera oblea probada pase y la segunda oblea probada falle, denotada como(*pasa, falla*), es

$$P[(pasa, falla)] = 0.8(0.2) = 0.16$$

La variable aleatoria X se define como el número de obleas que pasan. La última columna de la tabla muestra los valores de X que se asignan a cada resultado del experimento.

EJEMPLO 4-3

La variable aleatoria X se define como el número de partículas de contaminación en una oblea en la fabricación de semiconductores. Aun cuando las obleas poseen varias características, la variable aleatoria X resume las obleas tan sólo en términos del número de partículas.

Los valores posibles de X son enteros de cero hasta algún valor grande que representa el número máximo de partículas que pueden encontrarse en una de las obleas. Si este número máximo es muy grande, simplemente podría suponerse que el rango de X es el conjunto de los enteros desde cero hasta infinito.

Obsérvese que es posible definir más de una variable aleatoria en un espacio muestral. En el ejemplo 4-3, podría haberse definido la variable aleatoria Y como el número de chips de una oblea que no pasan la prueba final.

EJERCICIOS DE LA SECCIÓN 4-1

En cada uno de los ejercicios siguientes, determine el rango (los valores posibles) de la variable aleatoria.

4-1. La variable aleatoria es el número de conexiones soldadas que no cumplen con las especificaciones en una tarjeta de circuito impreso con 1 000 conexiones.

4-2. En un sistema de comunicación de voz con 50 líneas, la variable aleatoria es el número de líneas en uso en un tiempo dado.

4-3. Una báscula electrónica que muestra los pesos con la libra más próxima se usa para pesar paquetes. La pantalla sólo muestra cin-

co dígitos. Cualquier peso mayor del que puede mostrarse en la pantalla aparece como "99999." La variable aleatoria es el peso mostrado en la pantalla.

4-4. Un lote de 500 piezas maquinadas contiene 10 que no cumplen con los requerimientos del cliente. La variable aleatoria es el número de piezas, en una muestra de 5 piezas, que no cumplen con los requerimientos del cliente.

4-5. Un lote de 500 piezas maquinadas contiene 10 que no cumplen con los requerimientos del cliente. Se seleccionan piezas sucesivamente, sin reemplazo, hasta que se obtiene una pieza que no cumple. La variable aleatoria es el número de piezas seleccionadas.

4-6. La variable aleatoria es el contenido de humedad de un lote de materia prima, medido hasta el punto porcentual más próximo.

4-7. La variable aleatoria es el número de imperfecciones superficiales en un serpentín grande de acero galvanizado.

4-8. La variable aleatoria es el número de ciclos de un reloj de computadora requeridos para realizar un cálculo aritmético seleccionado.

4-9. En el pedido de un automóvil puede seleccionarse el modelo básico o agregar cualquier número de 15 opciones. La variable aleatoria es el número de opciones seleccionadas en un pedido.

4-10. Un pedido de entrepaños de madera puede hacerse con espesores de 1/8, 1/4 o 3/8 de pulgada. La variable aleatoria es el espesor total de los entrepaños de dos pedidos.

4-2 DISTRIBUCIONES DE PROBABILIDAD Y FUNCIONES DE MASA DE PROBABILIDAD

Muchas veces el interés se centra en la probabilidad de que una variable aleatoria asuma un valor particular.

La **distribución de probabilidad** de una variable aleatoria X es una descripción de las probabilidades asociadas con los valores posibles de X. Para una variable aleatoria discreta, es común especificar la distribución con una lista de los valores posibles junto con la probabilidad de cada uno. En algunos casos, resulta conveniente expresar la probabilidad en términos de una fórmula.

EJEMPLO 4-4

Existe la posibilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error. Sea X igual al número de bits con error en los cuatro siguientes bits transmitidos. Los valores posibles de X son $\{0, 1, 2, 3, 4\}$. Con base en un modelo para los errores que se presenta en la siguiente sección, se determinarán las probabilidades de estos valores. Suponga que las probabilidades son

$$P(X = 0) = 0.6561$$
$$P(X = 1) = 0.2916$$
$$P(X = 2) = 0.0486$$
$$P(X = 3) = 0.0036$$
$$P(X = 4) = 0.0001$$

La distribución de probabilidad de X se especifica por los valores posibles junto con la probabilidad de cada uno. En la figura 4-1 se presenta una descripción gráfica de la distribución de probabilidad de X.

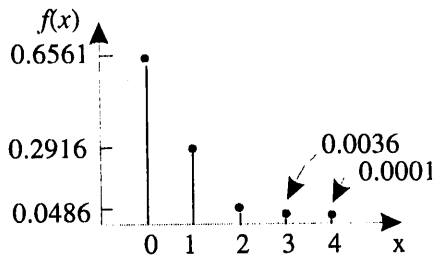


Figura 4-1 Distribución de probabilidad de los bits con error.

Suponga que una carga sobre una viga larga y delgada sólo coloca masa en puntos discretos de la misma (véase la figura 4-2). La carga puede describirse por una función que especifique la masa en cada uno de los puntos discretos. De manera similar, para una variable aleatoria discreta X, su distribución puede describirse por una función que especifique la probabilidad en cada uno de los valores discretos posibles de X.

Definición

Para una variable aleatoria discreta X con valores posibles $x_1, x_2, \ldots, x_n$, la **función de masa de probabilidad** es

$$f(x_i) = P(X = x_i). \tag{4-1}$$

Puesto que $f(x_i)$ se define como una probabilidad, $f(x_i) \geq 0$ para toda x_i y $\sum_{i=1}^{n} f(x_i) = 1$. Compruebe que la suma de las probabilidades del ejemplo 4-4 es uno.

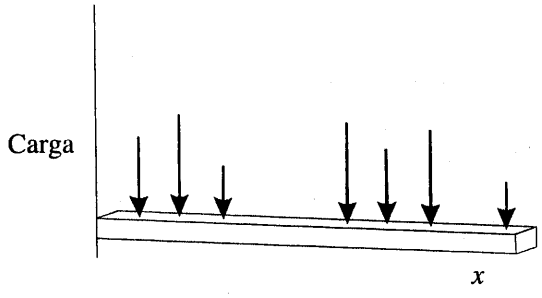


Figura 4-2 Cargas en puntos discretos sobre una viga larga y delgada.

EJEMPLO 4-5

Sea que la variable aleatoria X denote el número de obleas de semiconductores que es necesario analizar a fin de detectar una partícula grande de contaminación. Supóngase que la probabilidad de que una oblea contenga una partícula grande es 0.01 y que las obleas son independientes. Determine la distribución de probabilidad de X.

Sea que p denote una oblea en la que está presente una partícula grande y sea que a denote una oblea en la que está ausente. El espacio muestral del experimento es infinito y puede representarse como todas las secuencias posibles que empiecen con una cadena de *aes* y termine con una p. Es decir,

$$s = \{p,\ ap,\ aap,\ aaap,\ aaaap,\ aaaaap, \text{ y así sucesivamente}\}$$

Consideremos algunos casos especiales. Se tiene que $P(X = 1) = P(p) = 0.01$. Además, al emplear el supuesto de independencia,

$$P(X = 2) = P(ap) = 0.99(0.01) = 0.0099$$

Una fórmula general es

$$P(X = x) = \underbrace{P(aa \ldots ap)}_{(x-1)aes} = 0.99^{x-1}(0.01), \quad \text{para } x = 1, 2, 3, \ldots$$

Describir las probabilidades asociadas con X en términos de esta fórmula, es el método más sencillo para describir la distribución de X en este ejemplo.

Evidentemente, $f(x) \geq 0$. El hecho de que la suma de las probabilidades sea uno se deja como ejercicio.

Las variables aleatorias son a tal punto importantes en los experimentos aleatorios que en ocasiones se ignora de hecho el espacio muestral original del experimento y uno se concentra en la distribución de probabilidad de la variable aleatoria. Por ejemplo, en el ejemplo 4-1, el análisis podría enfocarse exclusivamente en los enteros $\{0, 1, \ldots, 48\}$ en el rango de X. En el ejemplo 4-2, el experimento aleatorio podría resumirse en términos de los tres valores posibles de X, es decir $\{0, 1, 2\}$. De esta manera, una variable aleatoria puede simplificar la descripción y el análisis de un experimento aleatorio.

EJERCICIOS DE LA SECCIÓN 4-2

4-11. El espacio muestral de un experimento aleatorio es $\{a, b, c, d, e, f\}$ y cada resultado es igualmente probable. Una variable aleatoria se define de la siguiente manera:

resultado	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

Determine la función de masa de probabilidad de X.

4-12. Utilice la función de masa de probabilidad del ejercicio 4-11 para determinar las siguientes probabilidades:

a) $P(X = 1.5)$

b) $P(0.5 < X < 2.7)$

c) $P(X > 3)$
d) $P(0 \leq X < 2)$
e) $P(X = 0 \text{ o } X = 2)$

Verifique que las siguientes funciones son funciones de masa de probabilidad y determine las probabilidades que se piden.

4-13.

x	-2	-1	0	1	2
$f(x)$	1/8	2/8	2/8	2/8	1/8

a) $P(X \leq 2)$
b) $P(X > -2)$
c) $P(-1 \leq X < 1)$
d) $P(X \leq -1 \text{ o } X = 2)$

4-14. $f(x) = (8/7)(1/2)^x$, X = 1, 2, 3
a) $P(X \leq 1)$
b) $P(X > 1)$
c) $P(2 < X < 6)$
d) $P(X \leq 1 \text{ o } X > 1)$

4-15. $f(x) = \dfrac{2x+1}{25}$, $x = 0, 1, 2, 3, 4$
a) $P(X = 4)$
b) $P(X \leq 1)$
c) $P(2 \leq X < 4)$
d) $P(X > -10)$

4-16. $f(x) = (3/4)(1/4)^x$, $x = 0, 1, 2, \ldots$
a) $P(X = 2)$
b) $P(X \leq 2)$
c) $P(X > 2)$
d) $P(X \geq 1)$

4-17. Estudios de mercado estiman que un nuevo instrumento para el análisis de muestras de suelo será de gran éxito, con un éxito moderado, o sin éxito, con probabilidades 0.3, 0.6 y 0.1, respectivamente. Los ingresos anuales asociados con un producto de gran éxito, con un éxito moderado o sin éxito son de \$10 millones, \$5 millones y \$1 millón, respectivamente. Sea que la variable aleatoria X denote los ingresos anuales del producto. Determinar la función de masa de probabilidad de X.

4-18. Un sistema de inspección óptica debe distinguir entre diferentes tipos de piezas. La probabilidad de una clasificación correcta de cualquier pieza es 0.98. Suponga que se inspeccionan tres piezas y que las clasificaciones son independientes. Sea que la variable aleatoria X denote el número de piezas que se clasifican correctamente. Determine la función de masa de probabilidad de X.

4-19. En un proceso de fabricación de semiconductores se prueban tres obleas de un lote. Cada oblea se clasifica como *pasa* o *falla*. Suponga que la probabilidad de que una oblea pase la prueba es 0.8 y que las obleas son independientes.
a) ¿Cuál es la probabilidad de que las tres obleas pasen la prueba?
b) Determine la función de masa de probabilidad del número de obleas de un lote que pasan la prueba.

4-20. Un ensamblaje consta de dos componentes mecánicos. Suponga que las probabilidades de que el primero y el segundo componentes cumplan con las especificaciones son 0.95 y 0.98. Suponga que los componentes son independientes.
a) ¿Cuál es la probabilidad de que todos los componentes de un ensamblaje cumplan con las especificaciones?
b) Determine la función de masa de probabilidad del número de componentes del ensamblaje que cumplen con las especificaciones.

4-21. Un ensamblaje consta de tres componentes mecánicos. Suponga que las probabilidades de que el primero, el segundo y el tercer componentes cumplan con las especificaciones son 0.95, 0.98 y 0.99. Suponga que los componentes son independientes.
a) ¿Cuál es la probabilidad de que todos los componentes de un ensamblaje cumplan con las especificaciones?
b) Determine la función de masa de probabilidad del número de componentes del ensamblaje que cumplen con las especificaciones.

4-22. Demuestre que la función $f(x)$ del ejemplo 4-5 satisface las propiedades de una función de masa de probabilidad haciendo la suma de la serie infinita.

4-3 FUNCIONES DE DISTRIBUCIÓN ACUMULADA

EJEMPLO 4-6

En el ejemplo 4-5, uno podría interesarse en la probabilidad de encontrar una partícula grande en tres o menos obleas. Este caso puede expresarse como $P(X \leq 3)$.

El evento de que $\{X \leq 3\}$ es la unión de los tres eventos $\{X = 1\}$, $\{X = 2\}$ y $\{X = 3\}$. El evento de que $\{X = 1\}$ consta del resultado único p, el evento de que $\{X = 2\}$ consta del resultado único ap y el evento de que $\{X = 3\}$ consta del resultado único aap. Evidentemente, estos tres eventos son mutuamente excluyentes. Por lo tanto,

$$\begin{aligned} P(X \leq 3) &= P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.01 + 0.99(0.01) + 0.99^2(0.01) \\ &= 0.0297 \end{aligned}$$

Para este ejemplo, puede presentarse una fórmula general,

$$P(X \leq x) = 1 - 0.99^x$$

Esta fórmula también puede usarse para determinar

$$\begin{aligned} P(X = 14) &= P(X \leq 14) - P(X \leq 13) \\ &= (1 - 0.99^{14}) - (1 - 0.99^{13}) \\ &= 0.99^{13} - 0.99^{14} \\ &= 0.99^{13}(0.01) \end{aligned}$$

El ejemplo 4-6 muestra que en ocasiones resulta conveniente poder expresar probabilidades acumuladas tales como $P(X \leq x)$ en términos de una fórmula y que una fórmula de las probabilidades acumuladas puede usarse para encontrar la función de masa de probabilidad de una variable aleatoria. Por lo tanto, utilizar probabilidades acumuladas es un método alternativo para describir la distribución de probabilidad de una variable aleatoria.

En general, para cualquier variable aleatoria discreta con valores posibles $x_1, x_2, \ldots, x_n$, los eventos $\{X = x_1\}$, $\{X = x_2\}$, ..., $\{X = x_n\}$ son mutuamente excluyentes. Por lo tanto, $P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.

Definición

La **función de distribución acumulada** de una variable aleatoria discreta X, denotada como $F(x)$, es

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

Para una variable aleatoria discreta X, $F(x)$ satisface las siguientes propiedades:

1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$
2) $0 \leq F(x) \leq 1$
3) Si $x \leq y$, entonces $F(x) \leq F(y)$ (4-2)

Al igual que una función de masa de probabilidad, una función de distribución acumulada proporciona probabilidades. Obsérvese que aun cuando la variable aleatoria X sólo puede asumir valores enteros, la función de distribución acumulada puede definirse en valores no enteros. En el ejemplo 4-6, $F(1.5) = P(X \leq 1.5) = P(X = 1) = 0.01$. Las propiedades 1 y 2 de la función de distribución acumulada se siguen de la definición. La propiedad 3 se sigue del hecho de que si $x \leq y$, entonces el evento de que $\{X \leq x\}$ está contenido en el evento $\{X \leq y\}$.

En el ejemplo siguiente se muestra cómo puede usarse la función de distribución acumulada para determinar la función de masa de probabilidad de una variable aleatoria discreta.

EJEMPLO 4-7

Suponga que la función de distribución acumulada de la variable aleatoria X es

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \leq x < 0 \\ 0.7 & 0 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

Determine la función de masa de probabilidad de X.

En la figura 4-3 se presenta una gráfica de $F(x)$. Por la gráfica, los únicos puntos que reciben una probabilidad diferente de cero son -2, 0 y 2 y

$$f(-2) = 0.2 - 0 = 0.2$$
$$f(0) = 0.7 - 0.2 = 0.5$$

y

$$f(2) = 1.0 - 0.7 = 0.3$$

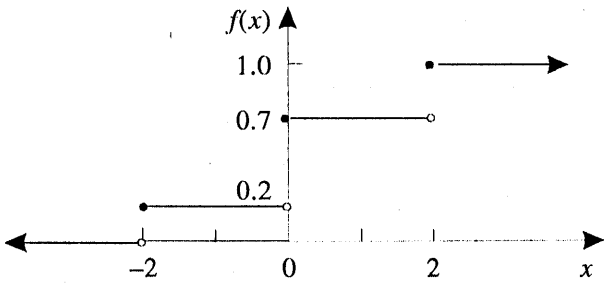


Figura 4-3 Función de distribución acumulada para el ejemplo 4-7.

EJEMPLO 4-8

Supóngase que la producción de un día de 850 piezas manufacturadas contiene 50 piezas que no cumplen con los requerimientos del cliente. Se seleccionan del lote dos piezas al azar y sin reemplazo. Sea la variable aleatoria X igual al número de piezas de la muestra que no cumplen. ¿Cuál es la función de distribución acumulada de X?

La pregunta puede contestarse encontrando primero la función de masa de probabilidad de X.

$$P(X = 0) = (800/850)(799/849) = 0.886$$
$$P(X = 1) = 2(800/850)(50/849) = 0.111$$
$$P(X = 2) = (50/850)(49/849) = 0.003$$

Por lo tanto,

$$F(0) = P(X \leq 0) = 0.886$$
$$F(1) = P(X \leq 1) = 0.0886 + 0.111 = 0.997$$
$$F(2) = P(X \leq 2) = 1$$

La función de distribución acumulada para este ejemplo se grafica en la figura 4-4. Obsérvese que $F(x)$ está definida para toda x de $-\infty < x < \infty$, no sólo para 0, 1 y 2.

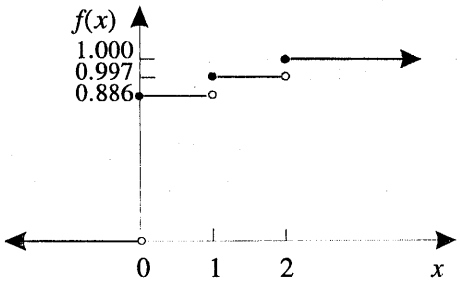


Figura 4-4 Función de distribución acumulada para el ejemplo 4-8.

EJERCICIOS DE LA SECCIÓN 4-3

4-23. Determine la función de distribución acumulada de la variable aleatoria del ejercicio 4-11.

4-24. Determine la función de distribución acumulada de la variable aleatoria del ejercicio 4-13; determine también las siguientes probabilidades:
a) $P(X \leq 1.25)$
b) $P(X \leq 2.2)$
c) $P(-1.1 < X \leq 1)$
d) $P(X > 0)$

4-25. Determine la función de distribución acumulada de la variable aleatoria del ejercicio 4-15; determine también las siguientes probabilidades:
a) $P(X < 1.5)$
b) $P(X \leq 3)$
c) $P(X > 2)$
d) $P(1 < X \leq 2)$

4-26. Determine la función de distribución acumulada de la variable aleatoria del ejercicio 4-17.

4-27. Determine la función de distribución acumulada de la variable aleatoria del ejercicio 4-19.

Verifique que las siguientes funciones son funciones de distribución acumulada y determine la función de masa de probabilidad y las probabilidades que se piden.

4-28.

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.5 & 1 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

a) $P(X \leq 3)$
b) $P(X \leq 2)$
c) $P(1 \leq X \leq 2)$
d) $P(X > 2)$

4-29.

$$F(x) = \begin{cases} 0 & x < -1 \\ 0.25 & -0.1 \leq x < 0.3 \\ 0.75 & 0.3 \leq x < 0.5 \\ 1 & 0.5 \leq x \end{cases}$$

a) $P(X \leq 0.5)$
b) $P(X \leq 0.4)$
c) $P(0.4 \leq X \leq 0.6)$
d) $P(X < 0)$
e) $P(0 \leq X < 0.1)$
f) $P(-0.1 < X < 0.1)$

4-30. El espesor del entrepaño de madera (en pulgadas) que pide un cliente es una variable aleatoria con la siguiente función de distribución acumulada:

$$F(x) = \begin{cases} 0 & x < 1/8 \\ 0.2 & 1/8 \leq x < 1/4 \\ 0.9 & 1/4 \leq x < 3/8 \\ 1 & 3/8 \leq x \end{cases}$$

Determine las siguientes probabilidades:
a) $P(X \leq 1/8)$
b) $P(X \leq 1/4)$
c) $P(X \leq 5/16)$
d) $P(X > 1/4)$
e) $P(X \leq 1/2)$

4-4 MEDIA Y VARIANZA DE UNA VARIABLE ALEATORIA DISCRETA

Así como resulta conveniente resumir una muestra de datos con la media y la varianza, la distribución de probabilidad de X se resume con su media y su varianza. Recuérdese que para los datos muestrales $x_1, x_2, \ldots, x_n$, la media muestral puede escribirse

$$\bar{x} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n$$

Es decir, $\bar{x}$ utiliza ponderaciones iguales de $1/n$ como factor de cada valor medido x_i. La media de una variable aleatoria X usa el modelo de probabilidad para ponderar los valores posibles de X. La media o valor esperado de X, denotado como μ o $E(X)$, es

$$\mu = E(X) = \sum_x xf(x)$$

Recuérdese que $\bar{x}$ es el punto de equilibrio cuando se coloca un peso igual en el lugar de cada medición en una recta numérica. De manera similar, si $f(x)$ es la función de densidad de una carga sobre una viga larga y delgada, entonces $E(X)$ es el punto en que la viga está en equilibrio. Por consiguiente, $E(X)$ describe el "centro" de la distribución de X de manera similar al punto de equilibrio de una carga.

Para los datos muestrales $x_1, x_2, ..., x_n$, la varianza es un resumen del esparcimiento o dispersión en los datos. Es

$$s^2 = \frac{1}{n-1}(x_1 - \bar{x})^2 + \frac{1}{n-1}(x_2 - \bar{x})^2 + \cdots + \frac{1}{n-1}(x_n - \bar{x})^2$$

Es decir, s^2 usa ponderaciones iguales de $1/(n-1)$ como factor de cada desviación cuadrada $(x_i - \bar{x})^2$. Como ya se mencionó, las desviaciones calculadas a partir de $\bar{x}$ tienden a ser menores que las calculadas a partir de μ y la ponderación se ajusta de $1/n$ a $1/(n-1)$ para compensar este hecho.

La varianza de una variable aleatoria X es una medida de dispersión o esparcimiento de los valores posibles de X. La varianza de X, denotada como σ^2 o $V(X)$, es

$$\sigma^2 = V(X) = \sum_x (x-\mu)^2 f(x)$$

$V(X)$ usa la ponderación $f(x)$ como el factor de cada posible desviación cuadrada $(x-\mu)^2$.

Utilizando las propiedades del operador sigma y la definición de μ puede demostrarse que

$$\begin{aligned} V(X) &= \sum_x (x-\mu)^2 f(x) \\ &= \sum_x x^2 f(x) - 2\mu \sum_x xf(x) + \mu^2 \sum_x f(x) \\ &= \sum_x x^2 f(x) - 2\mu^2 + \mu^2 = \sum_x x^2 f(x) - \mu^2 \end{aligned}$$

lo que permite usar una fórmula alternativa para $V(x)$.

Definición

La **media** o **valor esperado** de la variable aleatoria discreta X, denotada como μ o $E(X)$, es

$$\mu = E(X) = \sum_x xf(x) \tag{4-3}$$

La **varianza** de X, denotada como σ^2 o $V(X)$, es

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

La **desviación estándar** de X es $\sigma = [V(X)]^{1/2}$.

EJEMPLO 4-9

En el ejemplo 4-4, existe la posibilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error. Sea X igual al número de bits con error en los cuatro siguientes bits transmitidos. Los valores posibles de X son $\{0, 1, 2, 3, 4\}$. Con base en un modelo para los errores que se presenta en la sección siguiente, se determinarán las probabilidades de estos valores. Suponga que las probabilidades son:

$P(X = 0) = 0.6561$ $P(X = 2) = 0.0486$ $P(X = 4) = 0.0001$
$P(X = 1) = 0.2916$ $P(X = 3) = 0.0036$

Ahora bien,

$$\begin{aligned}\mu &= E(X) + 0f(0) + 1f(1) + 2f(2) + 3f(3) + 4f(4)\\ &= 0(0.6561) + 1(0.2916) + 2(0.0486) + 3(0.0036) + 4(0.0001)\\ &= 0.4\end{aligned}$$

Aun cuando X nunca asume el valor 0.4, el promedio ponderado de los valores posibles es 0.4.

Para calcular $V(X)$ resulta conveniente una tabla.

x	$x - 0.4$	$(x - 0.4)^2$	$f(x)$	$f(x)(x - 0.4)^2$
0	–0.4	0.16	0.6561	0.104976
1	0.6	0.36	0.2916	0.104976
2	1.6	2.56	0.0486	0.124416
3	2.6	6.76	0.0036	0.024336
4	3.6	12.96	0.0001	0.001296

$$V(X) = \sigma^2 = \sum_{i=1}^{5} f(x_i)(x_i - 0.4)^2 = 0.36$$

EJEMPLO 4-10

Va a hacerse la comparación del diseño de dos productos nuevos con base en la ganancia potencial. El departamento de comercialización considera que la ganancia del Diseño A puede predecirse con bastante precisión como \$3 millones. La ganancia potencial del Diseño B es más difícil de determinar. El departamento de comercialización concluye que hay una probabilidad de 0.3 de que la ganancia del Diseño B sea de \$7 millones, pero hay una probabilidad de 0.7 de que la ganancia sea de sólo \$2 millones. ¿Por cuál diseño se inclinaría el lector?

Sea que X denote la ganancia del Diseño A. Debido a que no hay incertidumbre en la ganancia del diseño A, el modelo de la distribución de la variable aleatoria X puede establecerse como \$3 millones con probabilidad uno. Por lo tanto, $E(X) = \$3$ millones.

Sea que Y denote la ganancia del diseño B. El valor esperado de Y en millones es

$$E(Y) = \$7(0.3) + \$2(0.7) = \$3.5$$

Debido a que $E(Y)$ es mayor que $E(X)$, podríamos inclinarnos por el Diseño B. Sin embargo, la variabilidad del resultado del Diseño B es mayor. Es decir,

$$\begin{aligned}\sigma^2 &= (7 - 3.5)^2(0.3) + (2 - 3.5)^2(0.7)\\ &= 5.25 \text{ millones de pesos al cuadrado}\end{aligned}$$

Puesto que las unidades de las variables en este ejemplo están en millones de pesos y como la varianza de una variable aleatoria eleva al cuadrado las desviaciones de la media, las unidades de σ^2 están en millones de pesos al cuadrado. Estas unidades dificultan la interpretación.

Puesto que las unidades de la desviación estándar son las mismas que las de la variable aleatoria, resulta más sencillo interpretrar la desviación estándar σ. En este ejemplo, los resultados pueden resumirse como "la desviación promedio que tiene Y de su media es \$2.29 millones".

EJEMPLO 4-11

El número de mensajes por hora enviados en una red de computadoras tiene la siguiente distribución:

x = número de mensajes	10	11	12	13	14	15
$f(x)$	0.08	0.15	0.30	0.20	0.20	0.07

Determine la media y la desviación estándar del número de mensajes por hora enviados.

$$\begin{aligned}E(X) &= 10(0.08) + 11(0.15) + \ldots + 15(0.07) = 12.5\\ V(X) &= 10^2(0.08) + 11^2(0.15) + \ldots + 15^2(0.07) - 12.5^2 = 1.85\\ \sigma &= \sqrt{V(X)} = \sqrt{1.85} = 1.36\end{aligned}$$

EJERCICIOS DE LA SECCIÓN 4-4

4-31. Si el rango de X es el conjunto $\{0, 1, 2, 3, 4\}$ y $P(X = x) = 0.2$, determine la media y la varianza de la variable aleatoria.

4-32. Determine la media y la varianza de la variable aleatoria del ejercicio 4-11.

4-33. Determine la media y la varianza de la variable aleatoria del ejercicio 4-13.

4-34. Determine la media y la varianza de la variable aleatoria del ejercicio 4-15.

4-35. Determine la media y la varianza de la variable aleatoria del ejercicio 4-17.

4-36. Determine la media y la varianza de la variable aleatoria del ejercicio 4-19.

4-37. El rango de la variable aleatoria X es $[0, 1, 2, 3, x]$, donde no se conoce el valor de x. Si cada valor es igualmente factible y la media de X es 6, determine x.

4-5 DISTRIBUCIÓN DISCRETA UNIFORME

El modelo de los resultados de varios sistemas físicos puede establecerse mediante los mismos experimentos aleatorios y variables aleatorias o similares. La distribución de las variables aleatorias presentes en cada uno de estos sistemas comunes puede analizarse, y es posible usar los resultados de ese análisis en diferentes aplicaciones y ejemplos. En este capítulo se presenta el análisis de varios experimentos aleatorios y variables aleatorias discretas que surgen con frecuencia en las aplicaciones. Con frecuencia se omite la exposición del espacio muestral fundamental del experimento aleatorio y se describe directamente la distribución de una variable aleatoria particular.

La variable aleatoria discreta más simple es aquella que sólo asume un número finito de posibles valores, cada uno con la misma probabilidad. Una variable aleatoria X que asume cada uno de los valores $x_1, x_2, ..., x_n$, con la misma probabilidad $1/n$, con frecuencia es de interés.

Definición

Una variable aleatoria X es una **variable aleatoria discreta uniforme** si cada uno de los n valores de su rango, por ejemplo, $x_1, x_2, ..., x_n$, tiene la misma probabilidad. Entonces,

$$f(x_i) = 1/n$$

EJEMPLO 4-12

El primer dígito del número de serie de una pieza es igualmente factible que sea cualquiera de los dígitos del 0 al 9. Si se selecciona una pieza de un lote grande y X es el primer dígito del número de serie, entonces X tiene una distribución discreta uniforme con probabilidad 0.1 para cada valor en $R = \{0, 1, 2, ..., 9\}$. Es decir,

$$f(x) = 0.1$$

para cada valor en R. La función de masa de probabilidad de X se muestra en la figura 4-5.

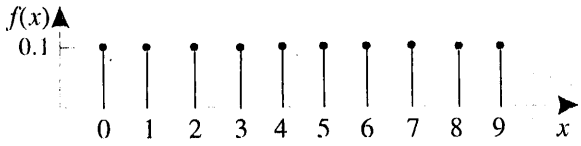


Figura 4-5 Función de masa de probabilidad para una variable aleatoria discreta uniforme.

Suponga que el rango de la variable aleatoria discreta X son los enteros consecutivos a, $a + 1$, $a + 2$, ..., b, para $a \leq b$. El rango de X contiene $b - a + 1$ valores, cada uno con probabilidad $1/(b - a + 1)$. Ahora bien,

$$\mu = \sum_{k=a}^{b} k\left(\frac{1}{b-a+1}\right) = \left(\frac{b(b+1)-(a-1)a}{2}\right)\left(\frac{1}{b-a+1}\right) = \frac{b+a}{2}$$

La deducción de la desviación estándar se deja como ejercicio.

Suponga que X es una variable aleatoria discreta uniforme en los enteros consecutivos a, $a + 1$, $a + 2$, ..., b, para $a \leq b$. La media de X es

$$\mu = E(X) = (b + a)/2$$

La desviación estándar de X es

$$\sigma = \sqrt{\frac{(b-a+1)^2 - 1}{12}} \tag{4-4}$$

EJEMPLO 4-13

En el ejemplo 4-1, sea que la variable aleatoria X denote el número de las 48 líneas de voz que están en uso en un tiempo dado. Suponga que X es una variable aleatoria discreta uniforme con rango de 0 a 48. Entonces,

$$E(X) = (48 + 0)/2 = 24$$

y

$$\sigma = \{[(48 - 0 + 1)^2 - 1]/12\}^{1/2} = 14.14$$

Las fórmulas de la ecuación 4-4 son más útiles de lo que pudiera parecer a primera vista. Si todos los valores en el rango de una variable aleatoria X se multiplican por una constante (sin

cambiar ninguna de las probabilidades), entonces la media y la desviación estándar de X se multiplican por la constante. En uno de los ejercicios se pide que el lector verifique este resultado. Puesto que la varianza de una variable aleatoria es el cuadrado de la desviación estándar, la varianza de X se multiplica por la constante al cuadrado. Resultados más generales de este tipo se exponen en el capítulo 6.

EJEMPLO 4-14

Sea que la variable aleatoria Y denote la proporción de las 48 líneas de voz que están en uso en un tiempo dado. Sea que X denote el número de líneas que están en uso en un tiempo dado. Entonces, $Y = X/48$. Por lo tanto,

$$E(Y) = E(X)/48 = 0.5$$

y

$$V(Y) = V(X)/48^2 = 0.087$$

EJERCICIOS DE LA SECCIÓN 4-5

4-38. Sea que la variable aleatoria X tenga una distribución discreta uniforme en los enteros $0 \leq x \leq 100$. Determine la media y la varianza de X.

4-39. Sea que la variable aleatoria X tenga una distribución discreta uniforme en los enteros $1 \leq x \leq 3$. Determine la media y la varianza de X.

4-40. Sea que la variable aleatoria X tenga la misma factibilidad de asumir cualquiera de los valores 1/8, 1/4 o 3/8. Determine la media y la varianza de X.

4-41. Se hacen mediciones del espesor de un proceso de recubrimiento hasta la centésima de milímetro más próxima. Las mediciones del espesor tienen una distribución uniforme con valores 0.15, 0.16, 0.17, 0.18 y 0.19. Determine la media y la varianza del espesor del recubrimiento para este proceso.

4-42. Se miden las longitudes de piezas de vidrio cilindrado hasta la décima de milímetro más próxima. Las longitudes tienen una distribución uniforme, con valores en cada décima de milímetro a partir de 590.0 y continuando hasta 590.9. Determine la media y la varianza de las longitudes.

4-43. Suponga que X tiene una distribución discreta uniforme en los enteros del 0 al 9. Determine la media, la varianza y la desviación estándar de la variable aleatoria $Y = 5X$ y haga la comparación con los resultados correspondientes para X.

4-44. Demuestre que, para una variable aleatoria discreta X, si cada uno de los valores en el rango de X se multiplica por la constante c, entonces el efecto es multiplicar la media de X por c y la varianza de X por c^2. Es decir, demuestre que $E(cX) = cE(X)$ y que $V(cX) = c^2V(X)$.

4-6 DISTRIBUCIÓN BINOMIAL

Considere los siguientes experimentos aleatorios y variables aleatorias:

1. El lanzamiento de una moneda 10 veces. Sea X = número de caras obtenidas.
2. Una máquina-herramienta desgastada produce 1% de piezas defectuosas. Sea X = número de piezas defectuosas en las siguientes 25 piezas producidas.

3. Cada muestra de aire tiene 10% de posibilidades de contener una molécula rara particular. Sea X = número de muestras de aire que contiene la molécula rara en las siguientes 18 muestras analizadas.
4. De todos los bits transmitidos a través de un canal de transmisión digital, 10% se reciben con error. Sea X = número de bits con error en los siguientes 5 bits transmitidos.
5. Un examen de opción múltiple contiene 10 preguntas, cada una con cuatro opciones, y todas las preguntas se contestan adivinando. Sea X = número de preguntas contestadas correctamente.
6. En los siguientes 20 nacimientos en un hospital, sea X = número de nacimientos de niñas.
7. De todos los pacientes que padecen una enfermedad particular, 35% experimentan una mejoría por un medicamento particular. En los siguientes 30 pacientes a los que se administra el medicamento, sea X = número de pacientes que experimentan una mejoría.

Estos ejemplos ilustran que un modelo general de probabilidad que los incluyera como casos particulares sería muy conveniente.

Puede considerarse que cada uno de estos experimentos aleatorios consta de una serie de ensayos aleatorios repetidos: 10 lanzamientos de la moneda en el experimento 1, la producción de 25 piezas en el experimento 2, etcétera. La variable aleatoria en cada caso es un conteo del número de ensayos que satisfacen el criterio especificado. El resultado de cada ensayo satisface el criterio para que X se cuente o bien no lo cuente; por consiguiente, el resultado de cada ensayo puede resumirse como éxito o bien como fracaso, respectivamente. Por ejemplo, en el experimento de opción múltiple únicamente se considera éxito la opción de cada pregunta que es la correcta. Elegir cualquiera de las tres opciones incorrectas en el ensayo se resume como un fracaso.

Los términos *éxito* y *fracaso* son meras etiquetas. Igual podría utilizarse "*A*" y "*B*" o "0" y "1". Desafortunadamente, las etiquetas comunes en ocasiones pueden ser engañosas. En el experimento 2, puesto que X es el conteo de las piezas defectuosas, a la producción de una pieza defectuosa se le llama éxito.

Un ensayo que sólo tiene dos resultados posibles se usa con tanta frecuencia como bloque de construcción de un experimento aleatorio que se conoce como **ensayo de Bernoulli**. Por lo general se supone que los ensayos que constituyen el experimento aleatorio son **independientes**. Esto implica que el resultado de un ensayo no afecta el resultado que se obtiene en cualquier otro ensayo. Además, con frecuencia es razonable suponer que la **probabilidad de un éxito en cada ensayo es constante**. En el experimento de opción múltiple, si la persona que toma la prueba no sabe nada del material y simplemente adivina en cada pregunta, podría suponerse que la probabilidad de una respuesta correcta es 1/4 *para cada pregunta.*

En esta sección se utiliza la notación factorial. Recuérdese que $n!$ denota el producto de los enteros menores o iguales que n

$$n! = n(n-1)(n-2)\ldots(2)(1)$$

Por ejemplo,

$$5! = (5)(4)(3)(2)(1) = 120$$
$$1! = 1$$

y por definición 0! = 1. También se usa la notación de combinaciones

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Por ejemplo,

$$\binom{5}{2} = \frac{5!}{2!\,3!} = \frac{120}{2 \cdot 6} = 10$$

Véase el apéndice I para comentarios adicionales.

EJEMPLO 4-15

Las posibilidades de que un bit transmitido a través de un canal de transmisión digital se reciba con error es 0.1. Suponga además que los ensayos de transmisión son independientes. Sea X = número de bits con error en los siguientes cuatro bits transmitidos. Determine $P(X = 2)$.

Sea que la letra E denote un bit con error y sea que la letra O denote que el bit está correcto, es decir, que se recibe sin error. Los resultados de este experimento pueden representarse como una lista de cuatro letras que indiquen los bits con error y los que están correctos. Por ejemplo, el resultado *OEOE* indica que el segundo y el cuarto bit tienen un error y que los dos bits restantes están correctos. Los valores correspondientes de x son:

Resultado	x	Resultado	x
OOOO	0	*EOOO*	1
OOOE	1	*EOOE*	2
OOEO	1	*EOEO*	2
OOEE	2	*EOEE*	3
OEOO	1	*EEOO*	2
OEOE	2	*EEOE*	3
OEEO	2	*EEEO*	3
OEEE	3	*EEEE*	4

El evento de que $X = 2$ consta de los seis resultados

$$\{EEOO, EOEO, EOOE, OEEO, OEOE, OOEE\}$$

Al utilizar el supuesto de que los ensayos son independientes, la probabilidad de $\{EEOO\}$ es

$$P(EEOO) = P(E)P(E)P(O)P(O) = (0.1)^2(0.9)^2 = 0.0081$$

además, cualquiera de los seis resultados mutuamente excluyentes para los que $X = 2$ tiene la misma probabilidad de ocurrir. Por lo tanto,

$$P(X = 2) = 6(0.0081) = 0.0486$$

En general,

$$P(X = x) = (\text{número de resultados que producen } x \text{ errores}) \text{ por } (0.1)^x(0.9)^{4-x}$$

Para completar una fórmula general de probabilidad sólo hace falta una expresión para el número de resultados que contienen x errores. Un resultado que contiene x errores puede construirse haciendo la partición de los cuatro ensayos (letras) del resultado en dos grupos. Un grupo es de tamaño x y contiene los errores, y el otro grupo es de tamaño $n - x$ y consta de los ensayos que están correctos. El número de formas de hacer la partición de cuatro objetos en dos grupos, uno de los cuales es de tamaño x, es $\binom{4}{x} = 4!/[x!(4-x)!]$. Por lo tanto, en este ejemplo,

$$P(X = x) = \binom{4}{x}(0.1)^x(0.9)^{4-x}$$

Obsérvese que $\binom{4}{2} = 4!/[2!\,2!] = 6$, como se encontró anteriormente. La función de masa de probabilidad de X se presentó en el ejemplo 4-4 y en la figura 4-1.

El ejemplo 4-15 motiva el resultado siguiente:

Definición

Un experimento aleatorio que consta de n ensayos repetidos tales que

1) los ensayos son independientes,
2) cada ensayo produce únicamente dos resultados posibles, etiquetados como "éxito" y "fracaso", y
3) la probabilidad de un éxito en cada ensayo, denotada como p, permanece constante

se llama *experimento binominal.*

La variable aleatoria X que es igual al número de ensayos que producen un éxito tiene una **distribución binomial** con parámetros p y $n = 1, 2, \ldots$.

La función de masa de probabilidad de X es

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n \tag{4-5}$$

Como en el ejemplo 4-15, $\binom{n}{x}$ es igual al número total de secuencias diferentes de ensayos que contienen x éxitos y $n - x$ fracasos. El número total de secuencias diferentes que contienen x éxitos y $n - x$ fracasos de veces la probabilidad de cada secuencia es igual a $P(X = x)$.

La expresión de probabilidad anterior es una fórmula muy útil que puede aplicarse en varios ejemplos. El nombre de la distribución se deriva de la *expansión binomial.* Para las constantes a y b, la expansión binomial es

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Sea que p denote la probabilidad de éxito en un solo ensayo. Entonces, utilizando la expansión binomial con $a = p$ y $b = 1 - p$, se observa que la suma de las probabilidades de una variable aleatoria binomial es uno. Además, debido a que cada ensayo del experimento se clasifica en dos resultados, {éxito, fracaso}, la distribución recibe el nombre de "bi"-nomial. Una distribución más general, la cual incluye la distribución binomial como un caso especial, es la distribución multinomial.

En la figura 4-6 se presentan ejemplos de distribuciones binomiales. Para un valor fijo de n, la distribución se hace cada vez más simétrica conforme p se incrementa de 0 a 0.5 o se decrementa de 1 a 0.5. Para un valor fijo de p, la distribución se hace cada vez más simétrica conforme n se incrementa.

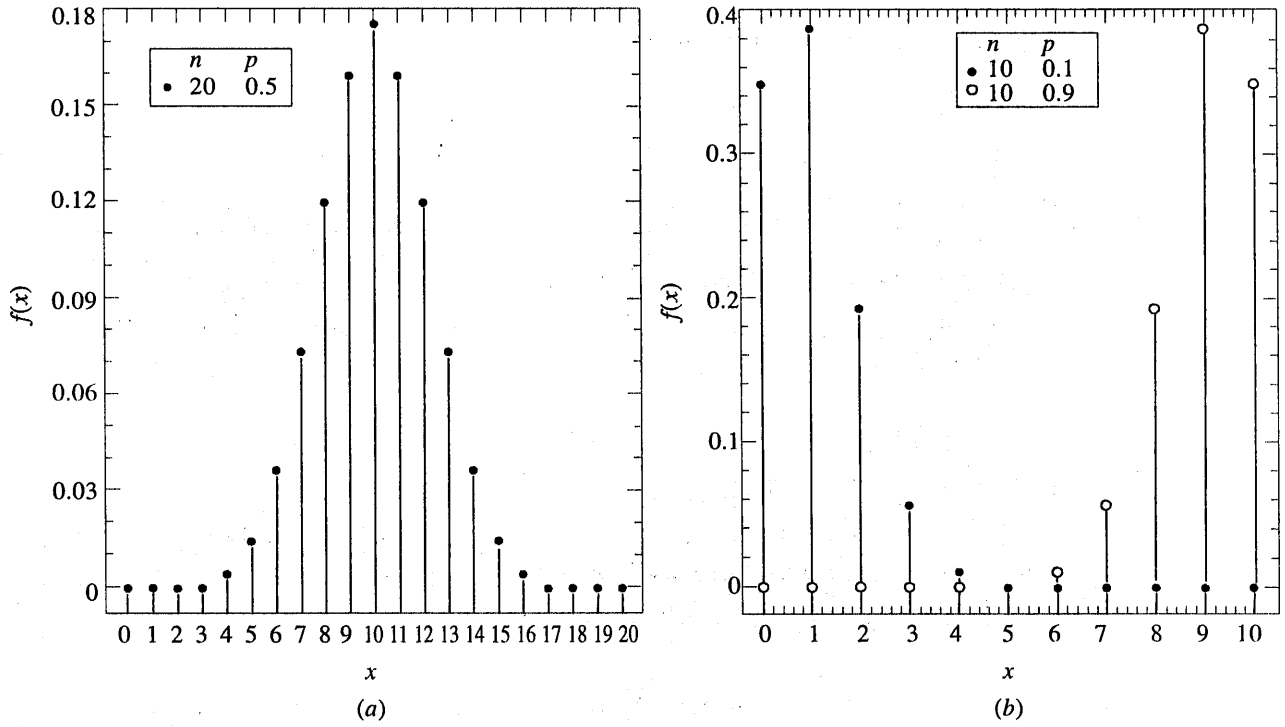


Figura 4-6 Distribuciones binominales para valores seleccionados de n y p.

EJEMPLO 4-16

Se presentan a continuación varios ejemplos en los que se emplea el coeficiente binomial $\binom{n}{x}$.

$$\binom{10}{3} = 10!/[3!\ 7!] = (10 \cdot 9 \cdot 8)(3 \cdot 2) = 120$$

$$\binom{15}{10} = 15!/[10!\ 5!] = (15 \cdot 14 \cdot 13 \cdot 12 \cdot 11)/(5 \cdot 4 \cdot 3 \cdot 2) = 3\ 003$$

$$\binom{100}{4} = 100!/[4!\ 96!] = (100 \cdot 99 \cdot 98 \cdot 97)/(4 \cdot 3 \cdot 2) = 3\ 921\ 225$$

EJEMPLO 4-17

Cada muestra de aire tiene 10% de posibilidades de contener una molécula rara particular. Suponga que las muestras son independientes con respecto a la presencia de la molécula rara. Encuentre la probabilidad de que en las siguientes 18 muestras, exactamente 2 contengan la molécula rara.

Sea X = número de muestras de aire que contienen la molécula rara en las siguientes 18 muestras analizadas. Entonces X es una variable aleatoria binomial con $p = 0.1$ y $n = 18$. Por lo tanto,

$$P(X=2)=\binom{18}{2}(0.1)^2(0.9)^{16}$$

Ahora bien, $\binom{18}{2} = (18!/[2!\ 16!]) = 18(17)/2 = 153$. Por lo tanto,

$$P(X=2)=153(0.1)^2(0.9)^{16}=0.284$$

Determine la probabilidad de que al menos cuatro muestras contengan la molécula rara. La probabilidad pedida es

$$P(X \geq 4)=\sum_{x=4}^{18}\binom{18}{x}(0.1)^x(0.9)^{18-x}$$

Sin embargo, es más sencillo usar el evento complementario,

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) \\ &= 1-\sum_{x=0}^{3}\binom{18}{x}(0.1)^x(0.9)^{18-x} \\ &= 1 - [0.150 + 0.300 + 0.284 + 0.168] \\ &= 0.098 \end{aligned}$$

Determine la probabilidad de que $3 \leq X < 7$. Ahora bien,

$$P(3 \leq X < 7) = \sum_{x=3}^{6} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$
$$= 0.168 + 0.070 + 0.022 + 0.005$$
$$= 0.265$$

La media y la varianza de una variable aleatoria binomial dependen únicamente de los parámetros p y n. Considérese el ejemplo de una máquina-herramienta desgastada para la que la probabilidad de una pieza defectuosa es p, y suponga que las piezas son independientes. Sea X = número de piezas defectuosas en las siguientes n piezas producidas. La media de X puede encontrarse considerando los n ensayos de Bernoulli del experimento. Se definen las nuevas variables aleatorias $X_1, ..., X_n$, tales que

$$X_k = \begin{cases} 1 & \text{si el ensayo } k \text{ es un éxito} \\ 0 & \text{en caso contrario} \end{cases}$$

Ahora bien,

$$X = \sum_{k=1}^{25} X_k$$

La media de cada X_k se obtiene de inmediato como

$$E(X_k) = 1p + 0(1 - p) = p$$

Entonces, la varianza de cada X_k es

$$V(X_k) = (1 - p)^2 p + (0 - p)^2 (1 - p)$$
$$= p(1 - p)$$

Puede demostrarse que

$$E\left(\sum_{k=1}^{25} X_k\right) = \sum_{k=1}^{25} E(X_k) = \sum_{k=1}^{25} p = 25p$$

y (dado que las X_k son independientes)

$$V\left(\sum_{k=1}^{25} X_k\right) = \sum_{k=1}^{25} V(X_k) = \sum_{k=1}^{25} p(1-p) = 25p(1-p)$$

Este ejemplo específico puede generalizarse en el siguiente resultado:

Si X es una variable aleatoria binomial con parámetros p y n, entonces

$$\mu = E(X) = np \quad \text{y} \quad \sigma^2 = V(X) = np(1-p) \tag{4-6}$$

EJEMPLO 4-18

Para el número de bits transmitidos que se reciben con error del ejemplo 4-15, $n = 4$ y $p = 0.1$, por lo que

$$E(X) = 4(0.1) = 0.4$$

La varianza de X es

$$V(X) = 4(0.1)(0.9) = 0.36$$

y estos resultados coinciden con los del ejemplo 4-9.

EJERCICIOS DE LA SECCIÓN 4-6

4-45. Para cada escenario que se describe a continuación, establezca si la distribución binomial es o no un modelo razonable para la variable aleatoria y por qué. Enuncie cualquier supuesto que se haga.

a) En un proceso de producción se fabrican miles de transductores de temperatura. Sea que X denote el número de transductores que no cumplen con las especificaciones en una muestra de tamaño 30 seleccionada al azar del proceso.

b) De un lote de 50 transductores de temperatura se selecciona una muestra de tamaño 30 sin reemplazo. Sea que X denote el número de transductores de esta muestra que no cumplen con las especificaciones.

c) Cuatro componentes electrónicos idénticos se conectan a un controlador que puede hacer el cambio de un componente inoperante a uno de los repuestos restantes. Sea que X denote el número de componentes que han presentado una falla después de un periodo de operación especificado.

d) Sea que X denote el número de accidentes que ocurren en las carreteras federales de Arizona durante un periodo de un mes.

e) Sea que X denote el número de respuestas correctas en un examen de opción múltiple en el que un estudiante puede eliminar algunas de las opciones cuando son incorrectas en algunas preguntas y todas las opciones incorrectas en otras preguntas.

f) Los defectos en la superficie de un chip semiconductor ocurren aleatoriamente. Sin embargo, sólo 80% de los defectos pueden detectarse mediante pruebas. Una muestra de 40 chips con un defecto cada uno se prueban. Sea que X denote el número de chips en los que la prueba detecta un defecto.

g) Reconsidere la situación del inciso f). Ahora, suponga que la muestra de 40 chips se compone de chips con uno y con cero defectos.

h) En una operación de llenado se intenta llenar paquetes de detergente con el peso publicitado. Sea que X denote el número de paquetes de detergente cuyo llenado no es completo.

i) Los errores en un canal de comunicación digital ocurren repentinamente, de tal manera que afectan a varios bits consecutivos. Sea que X denote el número de bits con error en una transmisión de 100 000 bits.

j) Sea que X denote el número de imperfecciones superficiales en un serpentín grande de acero galvanizado.

4-46. La variable aleatoria X tiene una distribución binomial con $n = 10$ y $p = 0.5$. Trace la función de masa de probabilidad de X.

a) ¿Cuál es el valor de X más factible?

b) ¿Cuál(es) es(son) el(los) valor(es) menos factible(s) de X?

4-47. La variable aleatoria X tiene una distribución binomial con $n = 10$ y $p = 0.5$. Determine las siguientes probabilidades:

a) $P(X = 5)$

b) $P(X \leq 2)$

c) $P(X \geq 9)$

d) $P(3 \leq X < 5)$

4-48. Trace la función de masa de probabilidad de una distribución binomial con $n = 10$ y $p = 0.01$ y comente el sesgo de la distribución.

a) ¿Cuál es el valor más factible de X?

b) ¿Cuál es el valor menos factible de X?

4-49. La variable aleatoria X tiene una distribución binomial con $n = 10$ y $p = 0.01$. Determine las siguientes probabilidades:

a) $P(X = 5)$

b) $P(X \leq 2)$

c) $P(X \geq 9)$

d) $P(3 \leq X < 5)$

4-50. Determine la función de distribución acumulada de una variable aleatoria binomial con $n = 3$ y $p = 1/2$.

4-51. Determine la función de distribución acumulada de una variable aleatoria binomial con $n = 3$ y $p = 1/4$.

4-52. Un producto electrónico contiene 40 circuitos integrados. La probabilidad de que cualquiera de los circuitos integrados esté defectuoso es 0.01, y los circuitos integrados son independientes. El producto solamente funciona si no hay circuitos integrados defectuosos. ¿Cuál es la probabilidad de que el producto funcione?

4-53. Sea que X denote el número de bits recibidos con error en un canal de comunicación digital, y suponga que X es una variable aleatoria binomial con $p = 0.001$. Si se transmiten 1 000 bits, determine lo siguiente:

a) $P(X = 1)$

b) $P(X \geq 1)$

c) $P(X \leq 2)$

d) La media y la varianza de X.

4-54. Las líneas telefónicas del sistema de reservaciones de una aerolínea están ocupadas 40% del tiempo. Suponga que los eventos de que las líneas estén ocupadas en llamadas sucesivas son independientes. Suponga que entran 10 llamadas a la aerolínea.

a) ¿Cuál es la probabilidad de que las líneas estén ocupadas para exactamente tres llamadas?

b) ¿Cuál es la probabilidad de que las líneas no estén ocupadas para al menos una llamada?

c) ¿Cuál es el número esperado de llamadas en las que todas las líneas estén ocupadas?

4-55. Se verifica si los lotes que constan de 50 resortes en espiral de un proceso de producción cumplen con los requerimientos del cliente. El número promedio de resortes en espiral de un lote que no cumplen con los requerimientos es 5. Suponga que el número de muelles que no cumplen en un lote, denotado como X, es una variable aleatoria binomial.

a) ¿Cuál es el valor de n y p?

b) ¿Cuál es $P(X \leq 2)$?

c) ¿Cuál es $P(X \geq 49)$?

4-56. Ejemplo de un diagrama de control estadístico de procesos. Cada hora se seleccionan muestras de 20 piezas de un proceso de punzón metálico. De manera típica, 1% de las piezas requieren reprocesamiento. Sea que X denote el número de piezas en la muestra de 20 que requieren reprocesamiento. Se sospecha un problema en el proceso, si X excede su media por más de tres desviaciones estándar.

a) Si el porcentaje de piezas que es necesario reprocesar se mantiene en 1%, ¿cuál es la probabilidad de que X exceda su media por más de tres desviaciones estándar?
b) Si el porcentaje del reprocesamiento aumenta a 4%, ¿cuál es la probabilidad de que X exceda el valor uno?
c) Si el porcentaje del reprocesamiento aumenta a 4%, ¿cuál es la probabilidad de que X exceda el valor uno en al menos una de las cinco siguientes horas de muestreo?

4-57. Debido a que no todos los pasajeros que hacen una reservación se presentan, una aerolínea vende 125 asientos para un vuelo con capacidad para sólo 120 pasajeros. La probabilidad de que un pasajero no se presente es 0.10 y el comportamiento de los pasajeros es independiente.
a) ¿Cuál es la probabilidad de que todos los pasajeros que se presenten puedan tomar el vuelo?
b) ¿Cuál es la probabilidad de que el vuelo parta con asientos vacíos?

4-58. Este ejercicio ilustra que las deficiencias en la calidad puede afectar programas y costos. En un proceso de manufactura tienen que surtirse los pedidos de 100 clientes. En cada pedido se necesita una parte componente que se compra a un proveedor. Sin embargo, de manera típica, 2% de los componentes se identifican como defectuosos, y puede suponerse que los componentes son independientes.
a) Si el fabricante tiene 100 componentes en existencia, ¿cuál es la probabilidad de que puedan surtirse los 100 pedidos sin reordenar componentes?
b) Si el fabricante tiene 102 componentes en existencia, ¿cuál es la probabilidad de que puedan surtirse los 100 pedidos sin reordenar componentes?
c) Si el fabricante tiene 105 componentes en existencia, ¿cuál es la probabilidad de que puedan surtirse los 100 pedidos sin reordenar componentes?

4-59. Un examen de opción múltiple contiene 25 preguntas con cuatro respuestas cada una. Suponga que un estudiante contesta cada una de las preguntas adivinando.
a) ¿Cuál es la probabilidad de que el estudiante responda correctamente más de 20 preguntas?
b) ¿Cuál es la probabilidad de que el estudiante responda correctamente menos de 5 preguntas?

4-60. Un semáforo particularmente tardado en el trayecto matutino del lector está en verde 20% de la veces que usted pasa por ese crucero. Suponga que cada mañana representa un ensayo independiente.
a) En 5 mañanas, ¿cuál es la probabilidad de que la luz esté en verde exactamente 1 día?
b) En 20 mañanas, ¿cuál es la probabilidad de que la luz esté en verde exactamente 4 días?
c) En 20 mañanas, ¿cuál es la probabilidad de que la luz esté en verde más de 4 días?

4-7 DISTRIBUCIONES GEOMÉTRICA Y BINOMIAL NEGATIVA

4-7.1 Distribución geométrica

Considérese un experimento aleatorio que guarda una estrecha relación con el que se utilizó en la definición de distribución binomial. De nueva cuenta, suponga una serie de ensayos de Bernoulli independientes con una probabilidad constante p de un éxito en cada ensayo. Sin embargo, en lugar de mantener fijo el número de ensayos, éstos se realizan hasta que se obtiene un éxito. Sea que la variable aleatoria X denote el número de ensayos hasta el primer éxito. En el ejemplo 4-5 se analizan obleas sucesivas hasta que se detecta una partícula grande. Entonces, X es el número de obleas analizadas. En la transmisión de bits, X podría ser el número de bits transmitidos hasta que ocurre un error.

EJEMPLO 4-19

La probabilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error es 0.1. Suponga que las transmisiones son eventos independientes y sea que la variable aleatoria X denote el número de bits transmitidos *hasta* el primer error.

Entonces, $P(X = 5)$ es la probabilidad de que los cuatro primeros bits se transmitan correctamente y que el quinto tenga un error. Este evento puede denotarse como $\{OOOOE\}$, donde O denota un bit correcto. Como los ensayos son independientes y la probabilidad de una transmisión correcta es 0.9,

$$P(X = 5) = P(OOOOE) = 0.9^4 0.1 = 0.066$$

Obsérvese que hay cierta probabilidad de que X sea igual a cualquier valor entero. Asimismo, si el primer ensayo es un éxito, entonces $X = 1$. Por lo tanto, el rango de X es $\{1, 2, 3, ...\}$, es decir, todos los enteros positivos.

Definición

En una serie de ensayos de Bernoulli independientes, con probabilidad constante p de un éxito, sea que la variable aleatoria X denote el número de ensayos hasta el primer éxito. Entonces X tiene una **distribución geométrica** con parámetro p y

$$f(x) = (1 - p)^{x-1}p, \quad x = 1, 2, ... \tag{4-7}$$

En la figura 4-7 se muestran ejemplos de las funciones de masa de probabilidad para variables aleatorias geométricas. Obsérvese que la altura de la línea en x es $(1 - p)$ veces la altura de la línea en $x - 1$. Es decir, las probabilidades decrecen en una progresión geométrica. La distribución recibe su nombre de este resultado.

EJEMPLO 4-20

La probabilidad de que una oblea contenga una partícula de contaminación grande es 0.01. Si se supone que las obleas son independientes, ¿cuál es la probabilidad de que sea necesario analizar exactamente 125 obleas antes de detectar una partícula grande?

Sea que X denote el número de muestras analizadas hasta que se detecta una partícula grande. Entonces X es una variable aleatoria geométrica con $p = 0.01$. La probabilidad pedida es

$$P(X = 125) = (0.99)^{124}0.01 = 0.0029$$

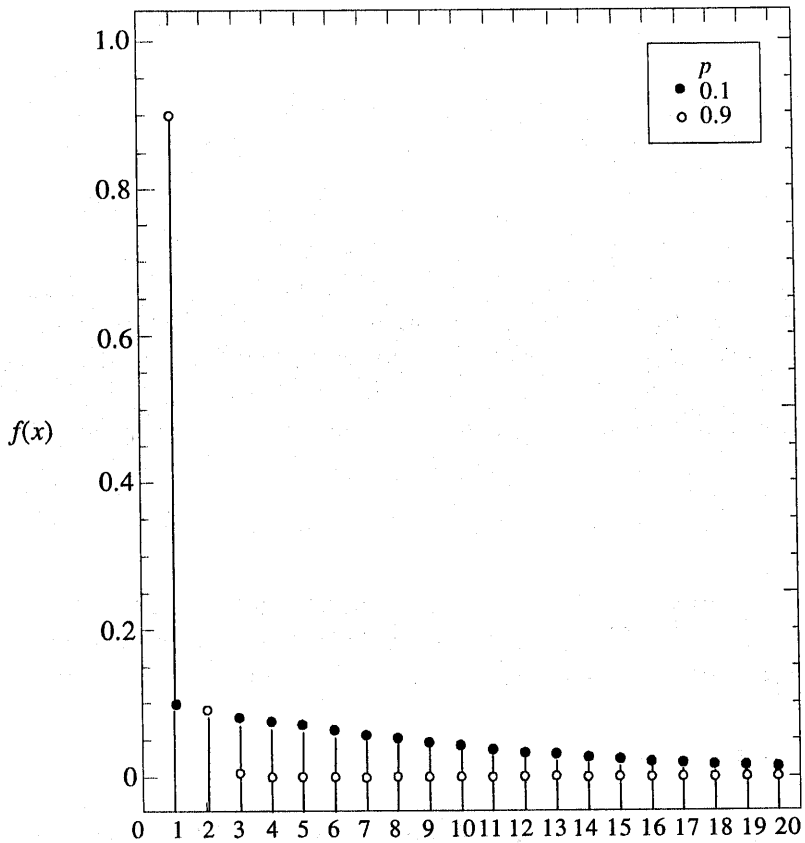


Figura 4-7 Distribuciones geométricas para valores seleccionados del parámetro p.

La deducción de la media y la varianza de una variable aleatoria geométrica se deja como ejercicio. Obsérvese que es posible demostrar que $\sum_{k=1}^{\infty} k\,(1-p)^{k-1}p$, es igual a $1/p$. Los resultados son los siguientes:

Si X es una variable aleatoria geométrica con parámetro p, entonces la media y la varianza de X son

$$\mu = E(X) = 1/p \quad \text{y} \quad \sigma^2 = V(X) = (1-p)/p^2 \tag{4-8}$$

EJEMPLO 4-21

Considere la transmisión de bits del ejemplo 4-19. Ahí, $p = 0.1$. El número promedio de transmisiones hasta el primer error es $1/0.1 = 10$. La desviación estándar del número de transmisiones antes del primer error es

$$\sigma = [(1-0.1)/0.1^2]^{1/2} = 9.49$$

Propiedad de falta de memoria

Una variable aleatoria geométrica se ha definido como el número de ensayos hasta el primer éxito. Sin embargo, puesto que los ensayos son independientes, el conteo del número de ensayos hasta el siguiente éxito puede iniciarse en cualquier ensayo sin cambiar la distribución de probabilidad de la variable aleatoria. Por ejemplo, en la transmisión de bits, si se transmiten 100 bits, la probabilidad de que el primer error, después del bit 100, ocurra en el bit 106 es la probabilidad de que los seis resultados siguientes sean *OOOOOE*. Esta probabilidad es $(0.9)^5(0.1) = 0.059$, que es idéntica a la probabilidad de que el primer error ocurra en el bit 6.

La consecuencia de utilizar un modelo geométrico es que el sistema probablemente nunca se agote. La probabilidad de un error se mantiene constante para todas las transmisiones. En este sentido, se dice que la distribución geométrica carece de memoria. La propiedad de falta de memoria se discutirá nuevamente en el contexto de una variable aleatoria exponencial en el capítulo 5.

EJEMPLO 4-22

En el ejemplo 4-19, la probabilidad de que un bit se transmita con error es igual a 0.1. Suponga que se han transmitido 50 bits. El número promedio de bits hasta el siguiente error es $1/0.1 = 10$, el mismo resultado que para el número promedio de bits hasta el primer error.

4-7.2 Distribución binomial negativa

Una generalización de una distribución geométrica en la que la variable aleatoria es el número de ensayos de Bernoulli requeridos para obtener r éxitos resulta en la **distribución binomial negativa**.

EJEMPLO 4-23

Como en el ejemplo 4-19, suponga que la probabilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error es 0.1. Suponga que las transmisiones son eventos independientes, y sea que la variable aleatoria X denote el número de bits transmitidos hasta el *cuarto* error.

Entonces, X tiene una distribución binomial negativa con $r = 4$. Las probabilidades que involucran a X pueden encontrarse como sigue: $P(X = 10)$ es la probabilidad de que ocurran exactamente tres errores en los nueve primeros ensayos y que el ensayo 10 sea el cuarto error. La probabilidad de que ocurran exactamente tres errores en los nueve primeros ensayos se determina a partir de la distribución binomial como

$$\binom{9}{3}(0.1)^3(0.9)^6$$

Puesto que los ensayos son independientes, la probabilidad de que ocurran exactamente tres errores en los nueve primeros ensayos y que el ensayo 10 resulte en el cuarto error es el producto de las probabilidades de estos dos eventos, es decir,

$$\binom{9}{3}(0.1)^3(0.9)^6(0.1) = \binom{9}{3}(0.1)^4(0.9)^6$$

El resultado anterior puede generalizarse como sigue:

Definición

En una serie de ensayos de Bernoulli independientes, con probabilidad constante p de éxito, sea que la variable aleatoria X denote el número de ensayos hasta que ocurran r éxitos. Entonces X tiene una **distribución binomial negativa** con parámetros p y $r = 1, 2, 3, \ldots$, y

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \tag{4-9}$$

para $x = r, r+1, r+2, \ldots$

Debido a que se necesitan al menos r ensayos para obtener r éxitos, el rango de X es de r a ∞. En el caso especial en que $r = 1$, una variable aleatoria binomial negativa es una variable aleatoria geométrica. En la figura 4-8 se ilustran algunas distribuciones binomiales negativas seleccionadas.

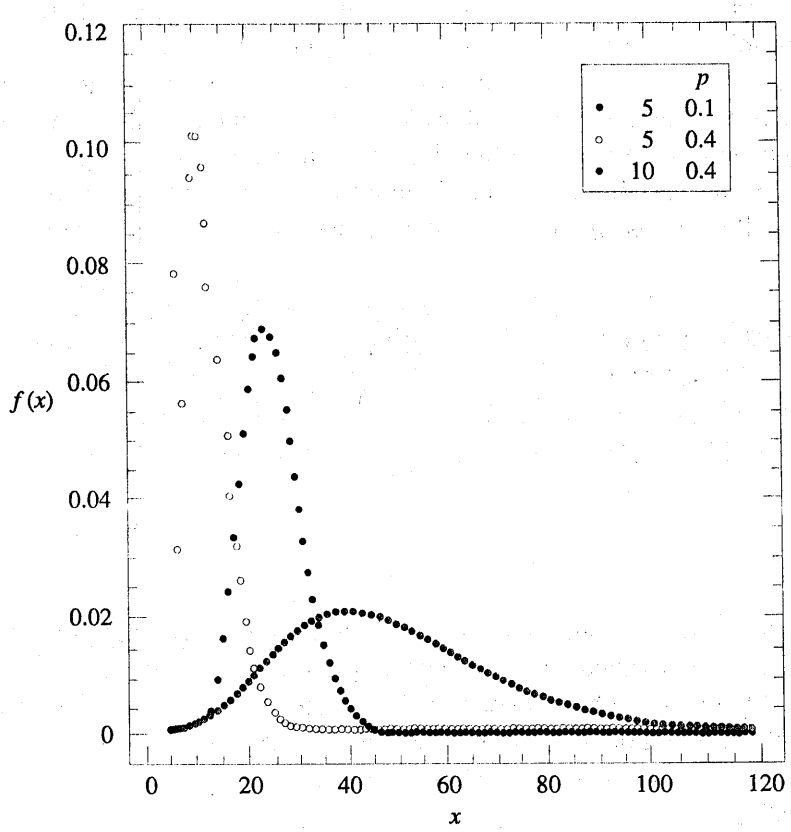
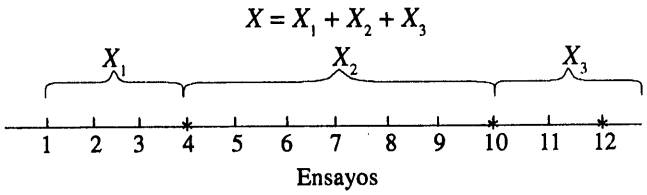


Figura 4-8 Distribuciones binomiales negativas para valores seleccionados de los parámetros r y p.



* indica un ensayo que resulta en un "éxito".

Figura 4-9 Una variable aleatoria binomial negativa representada como una suma de variables aleatorias geométricas.

La propiedad de falta de memoria de una variable aleatoria geométrica implica lo siguiente: sea que X denote el número de ensayos requeridos para obtener r éxitos, sea que X_1 denote el número de ensayos requeridos para obtener el primer éxito, sea que X_2 denote el número de ensayos adicionales requeridos para obtener el segundo éxito, sea que X_3 denote el número de ensayos adicionales requeridos para obtener el tercer éxito, y así sucesivamente. Entonces, el número total de ensayos requeridos para obtener r éxitos es $X = X_1 + X_2 + ... + X_r$. Debido a la propiedad de falta de memoria, cada una de las variables aleatorias $X_1, X_2, ..., X_r$ tiene una distribución geométrica con el mismo valor de p. Por consiguiente, una variable aleatoria binomial negativa puede interpretarse como la suma de r variables aleatorias geométricas. Este concepto se ilustra en la figura 4-9.

Recuérdese que una variable aleatoria binomial es un conteo del número de éxitos en n ensayos de Bernoulli. Es decir, el número de ensayos está predeterminado y el número de éxitos es aleatorio. Una variable aleatoria binomial negativa es un conteo del número de ensayos requeridos para obtener r éxitos. Es decir, el número de éxitos está predeterminado y el número de ensayos es aleatorio. En este sentido, una variable aleatoria binomial negativa puede interpretarse como el contrario, o negativo, de una variable aleatoria binomial.

La descripción de una variable aleatoria binomial negativa como una suma de variables aleatorias geométricas lleva a los siguientes resultados para la media y la varianza:

Si X es una variable aleatoria binomial negativa con parámetros p y r, entonces la media y la varianza de X son

$$\mu = E(X) = r/p \quad \text{y} \quad \sigma^2 = V(X) = r(1-p)/p^2 \tag{4-10}$$

EJEMPLO 4-24

Un avión de alto rendimiento contiene tres computadoras idénticas. Se utiliza únicamente una para operar el avión; las dos restantes son repuestos que pueden activarse en caso de que el sistema primario falle. Durante una hora de operación, la probabilidad de una falla en la computadora primaria (o de cualquiera de los sistemas de repuesto activados) es 0.0005. Suponiendo que cada hora representa un ensayo independiente, ¿cuál es el tiempo promedio para que fallen las tres computadoras?

Sea que X denote el número de horas hasta que los tres sistemas fallen, y sea que X_1, X_2 y X_3 denoten el número de horas de operación antes de una falla de la primera, la segunda y la tercera

computadoras usadas, respectivamente. Entonces, $X = X_1 + X_2 + X_3$. Además, se supone que las horas comprenden ensayos independientes con probabilidad constante de falla $p = 0.0005$. Por otra parte, una computadora de repuesto no es afectada por la cantidad de tiempo que transcurra antes de activarse. Por consiguiente, X tiene una distribución binomial negativa con $p = 0.0005$ y $r = 3$. En consecuencia,

$$E(X) = 3/0.0005 = 6\ 000 \text{ horas}$$

¿Cuál es la probabilidad de que las tres computadoras fallen en un vuelo de 5 horas? La probabilidad pedida es $P(X \leq 5)$ y

$$\begin{aligned} P(X \leq 5) &= P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.0005^3 + \binom{3}{2} 0.0005^3(0.9995) + \binom{4}{2} 0.0005^3(0.9995) \\ &= 1.25 \times 10^{-10} + 3.75 \times 10^{-10} + 7.49 \times 10^{-10} \\ &= 1.249 \times 10^{-9} \end{aligned}$$

EJERCICIOS DE LA SECCIÓN 4-7

4-61. Suponga que la variable aleatoria X tiene una distribución geométrica con $p = 0.5$. Determine las siguientes probabilidades:
a) $P(X = 1)$
b) $P(X = 4)$
c) $P(X = 8)$
d) $P(X \leq 2)$
e) $P(X > 2)$

4-62. Suponga que la variable aleatoria X tiene una distribución geométrica con una media de 2.5. Determine las siguientes probabilidades:
a) $P(X = 1)$
b) $P(X = 4)$
c) $P(X = 5)$
d) $P(X \leq 3)$
e) $P(X > 3)$

4-63. La probabilidad de una alineación óptica de éxito en el ensamblaje de un producto de almacenamiento óptico de datos es 0.8. Suponga que los ensayos son independientes.
a) ¿Cuál es la probabilidad de que la primera alineación de éxito requiera exactamente cuatro ensayos?
b) ¿Cuál es la probabilidad de que la primera alineación de éxito requiera a lo sumo cuatro ensayos?
c) ¿Cuál es la probabilidad de que la primera alineación de éxito requiera al menos cuatro ensayos?

4-64. Suponga que cada llamada telefónica a una popular radiodifusora tiene una probabilidad de 0.02 de conseguir una conexión, es decir, de no obtener una señal de ocupado. Suponga que las llamadas son independientes.
a) ¿Cuál es la probabilidad de que la primera llamada que consiga una conexión sea la décima?
b) ¿Cuál es la probabilidad de que se requieran más de cinco llamadas para conseguir una conexión?
c) ¿Cuál es el número promedio de las llamadas necesarias para conseguir una conexión?

4-65. En el ejercicio 4-60, recuérdese que un semáforo particularmente tardado en el trayecto matutino del lector está en verde 20% de la veces que usted pasa por ese crucero. Suponga que cada mañana representa un ensayo independiente.
a) ¿Cuál es la probabilidad de que la primera mañana que el semáforo esté en verde

sea la cuarta mañana que usted pasa por el crucero?

b) ¿Cuál es la probabilidad de que el semáforo no esté en verde en 10 mañanas consecutivas?

4-66. En el ejercicio 4-56, recuérdese que cada hora se seleccionan muestras de 20 piezas de un proceso de punzón metálico y que X denota el número de piezas de la muestra de 20 que requieren reprocesamiento.

a) Si el porcentaje de piezas que requieren reprocesamiento se mantiene en 1%, ¿cuál es la probabilidad de que la muestra de la hora 10 sea la primera en que X exceda el valor uno?

b) Si el porcentaje de reprocesamiento aumenta a 4%, ¿cuál es la probabilidad de que la muestra de la hora 10 sea la primera en que X exceda el valor uno?

c) Si el porcentaje de reprocesamiento aumenta a 4%, ¿cuál es el número esperado de horas hasta que X exceda el valor uno?

4-67. Considere una sucesión de ensayos de Bernoulli independientes con $p = 0.2$.

a) ¿Cuál es el número esperado de ensayos para obtener el primer éxito?

b) Después de ocurrir el octavo éxito, ¿cuál es el número esperado de ensayos para obtener el noveno éxito?

4-68. Demuestre que la función de densidad de probabilidad de una variable aleatoria binomial negativa es igual a la función de densidad de probabilidad de una variable aleatoria geométrica cuando $r = 1$. Demuestre que la fórmula de la media y la varianza de una variable aleatoria binomial negativa es igual a los resultados correspondientes para una variable aleatoria geométrica cuando $r = 1$.

4-69. Suponga que X es una variable aleatoria binomial negativa con $p = 0.2$ y $r = 4$. Determine lo siguiente:

a) $E(X)$

b) $P(X = 20)$

c) $P(X = 19)$

d) $P(X = 21)$

e) El valor más probable de X.

4-70. Es de 0.6 la probabilidad de que la calibración de un transductor en un instrumento electrónico cumpla con las especificaciones del sistema de medición. Suponga que las operaciones de calibración son independientes. ¿Cuál es la probabilidad de que se necesiten a lo sumo tres operaciones de calibración para cumplir con las especificaciones del sistema de medición?

4-71. Una báscula electrónica en una operación de llenado automatizado detiene la línea de producción después de detectar tres paquetes con peso insuficiente. Suponga que la probabilidad de un paquete con peso insuficiente es 0.001 y que cada operación de llenado es independiente.

a) ¿Cuál es el número promedio de operaciones de llenado antes de que se detenga la línea de producción?

b) ¿Cuál es la desviación estándar del número de operaciones de llenado antes de que se detenga la línea de producción?

4-72. Un sistema con tolerancia para fallas que procesa las transacciones de una empresa de servicios financieros utiliza tres computadoras separadas. Si la computadora en operación falla, una de las dos computadoras de repuesto puede ponerse en línea de inmediato. Después de que falla la segunda computadora, la última computadora puede ponerse en línea de inmediato. Suponga que la probabilidad de una falla durante cualquier transacción es 10^{-8} y que las transacciones pueden considerarse eventos independientes.

a) ¿Cuál es el número promedio de transacciones antes de que las tres computadoras hayan fallado?

b) ¿Cuál es la varianza del número de transacciones antes de que todas las computadoras hayan fallado?

4-73. Deduzca las expresiones para la media y la varianza de una variable aleatoria geométrica con parámetro p. (Se requieren fórmulas de series infinitas.)

4-8 DISTRIBUCIÓN HIPERGEOMÉTRICA

En el ejemplo 4-8, la producción de un día de 850 piezas manufacturadas contiene 50 piezas que no cumplen con los requerimientos del cliente. Se seleccionan dos piezas, al azar y sin reemplazo, de la producción de un día. ¿Cuál es la probabilidad de que en la muestra haya dos piezas que no cumplan los requerimientos?

Este experimento consta de dos ensayos: se selecciona la primera pieza y luego se selecciona la segunda. Sin embargo, este experimento presenta diferencias fundamentales con los ejemplos basados en la distribución binomial. En este experimento, los ensayos no son independientes. Obsérvese que en el inusual caso de que se hiciera el reemplazo de cada unidad seleccionada antes de la siguiente selección, los ensayos serían independientes y la probabilidad de que una pieza que no cumpla en cada ensayo sería constante. En tal caso, el número de piezas de la muestra que no cumplen, es una variable aleatoria binomial.

Sea que A y B denoten los eventos de que la primera y la segunda pieza no cumplen, respectivamente. En el capítulo 3 se encontró que $P(B \mid A) = 49/849$ y que $P(A) = 50/850$. Por consiguiente, saber que la primera pieza no cumple, sugiere que es menos factible que la segunda pieza seleccionada no cumpla con los requerimientos del cliente.

Sea X igual al número de piezas de la muestra que no cumplen. Entonces,

$P(X = 0) =$ P(ambas piezas cumplen) $= (800/850)(799/849) = 0.886$

$P(X = 1) =$ P(la primera pieza cumple y la segunda no, o la primera pieza seleccionada no cumple y la segunda pieza seleccionada sí)

$= (800/850)(50/849) + (50/850)(800/849) = 0.111$

$P(X = 2) =$ P(ambas partes no cumplen) $= (50/850)(49/849) = 0.003$

Como en este ejemplo, con frecuencia las muestras se seleccionan *sin reemplazo*; es decir, las unidades seleccionadas no se reemplazan antes de hacer la siguiente selección. Aun cuando las probabilidades pueden determinarse mediante el razonamiento empleado en el ejemplo anterior, es de gran utilidad una fórmula general para calcular las probabilidades cuando se seleccionan muestras sin reemplazo. Pueden utilizarse las reglas de conteo del apéndice I para justificar la fórmula que se presenta a continuación:

Definición

Un conjunto de N objetos contiene

K objetos clasificados como éxitos y

$N - K$ objetos clasificados como fracasos

Se selecciona una muestra con tamaño de n objetos al azar (sin reemplazo) de los N objetos, donde $K \leq N$ y $n \leq N$.

Sea que la variable aleatoria X denote el número de éxitos en la muestra. Entonces X tiene una **distribución hipergeométrica** y

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \qquad x = \text{máx}\{0, n + K - N\} \text{ hasta mín } \{K, n\} \tag{4-11}$$

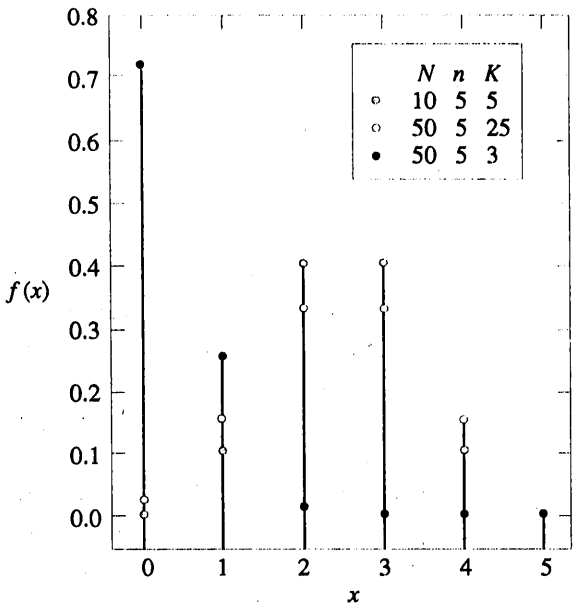


Figura 4-10 Distribuciones hipergeométricas para valores escogidos de los parámetros N, K y n.

Se usa la expresión mín$\{K, n\}$ en la definición del rango de X porque el número máximo de éxitos que pueden ocurrir en la muestra es el menor entre el tamaño de la muestra, n, y el número de éxitos disponibles, K. Asimismo, si $n + K > N$, deben ocurrir al menos $n + K - N$ éxitos en la muestra. En la figura 4-10 se ilustran algunas distribuciones hipergeométricas escogidas.

EJEMPLO 4-25

El ejemplo del principio de esta sección puede analizarse de nuevo utilizando la expresión general de la definición de variable aleatoria hipergeométrica. Es decir,

$$P(X=0)=\frac{\binom{50}{0}\binom{800}{2}}{\binom{850}{2}}=\frac{319\,600}{360\,825}=0.886$$

$$P(X=1)=\frac{\binom{50}{1}\binom{800}{1}}{\binom{850}{2}}=\frac{40\,000}{360\,825}=0.111$$

$$P(X=2)=\frac{\binom{50}{2}\binom{800}{0}}{\binom{850}{2}}=\frac{1\,225}{360\,825}=0.003$$

EJEMPLO 4-26

Un lote contiene 100 piezas de un proveedor de tubería local y 200 unidades de un proveedor de tubería del estado vecino. Si se seleccionan cuatro piezas al azar y sin reemplazo, ¿cuál es la probabilidad de que todas sean del proveedor local?

Sea X igual al número de piezas de la muestra del proveedor local. Entonces, X tiene una distribución hipergeométrica y la probabilidad pedida es $P(X = 4)$. Por consiguiente,

$$P(X=4)=\frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}=0.0119$$

¿Cuál es la probabilidad de que dos o más piezas de la muestra sean del proveedor local?

$$P(X\geq 2)=\frac{\binom{100}{2}\binom{200}{2}}{\binom{300}{4}}+\frac{\binom{100}{3}\binom{200}{1}}{\binom{300}{4}}+\frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}$$

$$=0.298+0.098+0.0119=0.408$$

¿Cuál es la probabilidad de que al menos una pieza de la muestra sea del proveedor local?

$$P(X\geq 1)=1-P(X=0)=1-\frac{\binom{100}{0}\binom{200}{4}}{\binom{300}{4}}=0.196$$

La media y la varianza de una variable aleatoria hipergeométrica pueden determinarse considerando los ensayos que componen el experimento. Sin embargo, los ensayos no son independientes y, por tanto, los cálculos que se usaron para una variable aleatoria binomial no son apropiados. La media y la varianza de una variable aleatoria hipergeométrica se dejan como ejercicio para desarrollar el intelecto.

Sea $p = K/N$. Entonces, p se interpreta como la proporción de éxitos en el conjunto del que se elige la muestra.

Si X es una variable aleatoria hipergeométrica con parámetros N, K y n, entonces la media y la varianza de X son

$$\mu=E(X)=np \quad y \quad \sigma^2=V(X)=np(1-p)\frac{N-n}{N-1} \tag{4-12}$$

donde $p = K/N$.

EJEMPLO 4-27

En el ejemplo anterior, el tamaño de la muestra es 4. La variable aleatoria X es el número de piezas de la muestra del proveedor local. Entonces, $p = 100/300 = 1/3$. Por lo tanto,

$$E(X) = 4(100/300) = 1.33$$

y

$$V(X) = 4(1/3)(2/3)[(300 - 4)/299] = 0.88$$

Obsérvese que para una variable aleatoria hipergeométrica, $E(X)$ es similar al resultado para una variable aleatoria binomial. Además, $V(X)$ difiere del resultado para una variable aleatoria binomial tan sólo por el término que se presenta en seguida:

La expresión

$$\frac{N - n}{N - 1}$$

se conoce como el **factor de corrección para poblaciones finitas.**

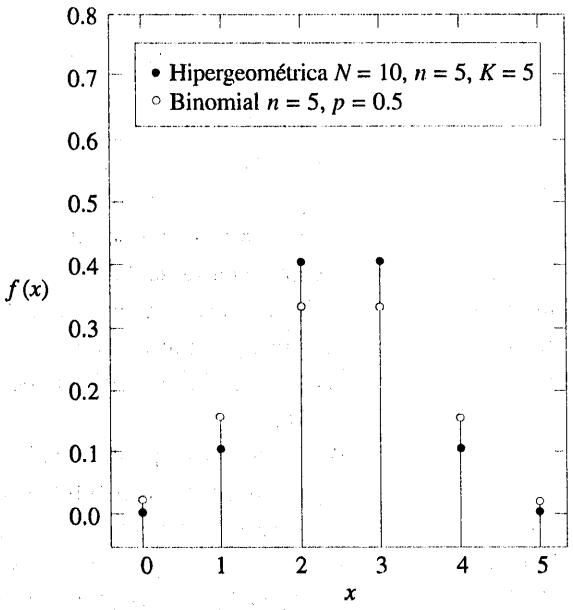
Hacer un muestreo con reemplazo es equivalente a hacer un muestreo de un conjunto infinito porque la proporción de éxitos se mantiene constante para cada ensayo del experimento. Como ya se mencionó, si el muestreo se hiciera con reemplazo, entonces X sería una variable aleatoria binomial y su varianza sería $np(1 - p)$. Por consiguiente, la corrección para poblaciones finitas representa la corrección de la varianza binomial que resulta porque el muestreo se hace sin reemplazo del conjunto finito de tamaño N.

Si n es pequeña en comparación con N, entonces la corrección es pequeña y la distribución hipergeométrica es similar a la binomial. En este caso, aun si el muestreo se hiciera *sin reemplazo,* una distribución binomial puede ser una aproximación efectiva de la distribución del número de unidades de la muestra de un tipo especificado. En la figura 4-11 se presenta un ejemplo.

EJEMPLO 4-28

La lista de cuentas de clientes de una corporación grande contiene 1 000 clientes. De ellos, 700 han comprado al menos uno de los productos de la corporación en los últimos tres meses. Para evaluar el diseño de un nuevo producto, se hace un muestreo aleatorio de 50 clientes de la lista de la corporación. ¿Cuál es la probabilidad de que más de 45 de los clientes muestreados hayan hecho una compra de la corporación en los últimos tres meses?

El muestreo es sin reemplazo. Sin embargo, debido a que el tamaño de la muestra de 50 es relativamente pequeño en comparación con el número de cuentas de clientes, 1 000, la probabilidad de seleccionar a un cliente que haya hecho una compra de la corporación en los últimos tres meses se mantiene relativamente constante conforme se elige a los clientes.



	0	1	2	3	4	5
Probabilidad hipergeométrica	0.004	0.099	0.397	0.397	0.099	0.004
Probabilidad binomial	0.031	0.156	0.313	0.313	0.156	0.031

Figura 4-11 Comparaciones de las distribuciones hipergeométrica y binomial para valores escogidos de los parámetros.

Por ejemplo, sea que A denote el evento de que el primer cliente seleccionado no ha hecho una compra de la corporación en los últimos tres meses y sea que B denote el evento de que el segundo cliente seleccionado no ha hecho una compra de la corporación en los últimos tres meses. Entonces $P(A) = 300/1\,000 = 0.3$ y $P(B \mid A) = 299/999 = 0.2993$. Esto es, los intentos son aproximadamente independientes. En la muestra, sea X la que denota el número de clientes que han hecho compras de la corporación en los últimos tres meses. Entonces, X es una variable aleatoria hipergeométrica con $N = 1\,000$, $n = 50$ y $K = 700$. Por consiguiente, $p = K/N = 0.7$. La probabilidad pedida es $P(X > 45)$. Debido a que el tamaño de la muestra es relativamente pequeño en comparación con el tamaño de la población, la distribución de X puede aproximarse considerándola binomial con $n = 50$ y $p = 0.7$. Al utilizar la aproximación binomial para la distribución de X se obtiene

$$P(X > 45) = \sum_{x=46}^{50} \binom{50}{x} 0.7^x (1-0.7)^{50-x} = 0.00017$$

EJERCICIOS DE LA SECCIÓN 4-8

4-74. Suponga que X tiene una distribución hipergeométrica con $N = 100$, $n = 4$ y $K = 20$. Determine lo siguiente:

a) $P(X = 1)$
b) $P(X = 6)$
c) $P(X = 4)$
d) Determine la media y la varianza de X.

4-75. Suponga que X tiene una distribución hipergeométrica con $N = 20$, $n = 4$ y $K = 4$. Determine lo siguiente:

a) $P(X = 1)$
b) $P(X = 4)$
c) $P(X \leq 2)$
d) Determine la media y la varianza de X.

4-76. Suponga que X tiene una distribución hipergeométrica con $N = 10$, $n = 3$ y $K = 4$. Trace la función de masa de probabilidad de X.

4-77. Determine la función de distribución acumulada de X en el ejercicio 4-76.

4-78. Un lote de 75 arandelas contiene cinco en las que la variabilidad del espesor alrededor de la circunferencia de la arandela es inaceptable. Se selecciona, al azar y sin reemplazo, una muestra de 10 arandelas.

a) ¿Cuál es la probabilidad de que ninguna de las arandelas inaceptables esté en la muestra?
b) ¿Cuál es la probabilidad de que al menos una de las arandelas inaceptables esté en la muestra?
c) ¿Cuál es la probabilidad de que exactamente una de las arandelas inaceptables esté en la muestra?
d) ¿Cuál es el número promedio de arandelas inaceptables en la muestra?

4-79. Tarjetas de circuitos impresos se someten a pruebas de funcionamiento después de instalar en ellas chips semiconductores. Un lote contiene 140 tarjetas, y se seleccionan 20 sin reemplazo para las pruebas de funcionamiento.

a) Si hay 20 tarjetas defectuosas, ¿cuál es la probabilidad de que al menos una de las tarjetas defectuosas esté en la muestra?
b) Si hay cinco tarjetas defectuosas, ¿cuál es la probabilidad de que al menos una de las tarjetas defectuosas aparezca en la muestra?

4-80. Una cinta magnética se corta en tiras de media pulgada de ancho que se enrollan en carretes. Un montaje de cuchillas contiene 48 navajas. Diariamente se seleccionan al azar cinco navajas y se evalúa su filo. Si se encuentra alguna navaja desafilada, el montaje se reemplaza con un juego nuevo de navajas afiladas.

a) Si hay 10 navajas desafiladas en un montaje, ¿cuál es la probabilidad de que el montaje se reemplace el primer día que se evalúa?
b) Si hay 10 navajas desafiladas en un montaje, ¿cuál es la probabilidad de que el montaje no se reemplace sino hasta el tercer día que se evalúa? [*Sugerencia:* Suponga que las decisiones diarias son independientes y use la distribución geométrica.]
c) Suponga que en el primer día de evaluación dos de las navajas están desafiladas, el segundo día de evaluación seis están desafiladas y el tercer día de evaluación 10 están desafiladas. ¿Cuál es la probabilidad de que el ensamblaje no se reemplace sino hasta el tercer día de evaluación? [*Sugerencia:* Suponga que las decisiones diarias son independientes. Sin embargo, la probabilidad de reemplazamiento cambia cada día.]

4-81. Un estado opera una lotería en la que se seleccionan aleatoriamente seis números de 40, sin reemplazo. Un jugador elige seis números antes de que se seleccione la muestra del estado.

a) ¿Cuál es la probabilidad de que los seis números elegidos por un jugador coincidan con los seis números de la muestra del estado?
b) ¿Cuál es la probabilidad de que cinco de los seis números elegidos por un jugador aparezcan en la muestra del estado?
c) ¿Cuál es la probabilidad de que cuatro de los seis números elegidos por un jugador aparezcan en la muestra del estado?

d) Si un jugador entra en una lotería cada semana, ¿cuál es el número esperado de semanas hasta que el jugador le atine a los seis números de la muestra del estado?

4-82. Continuación de los ejercicios 4-74 y 4-75.

a) Calcule la corrección para poblaciones finitas de los ejercicios 4-74 y 4-75. ¿Para cuál de los dos ejercicios resultaría más conveniente la aproximación binomial de la distribución de X?

b) Para el ejercicio 4-74, calcule $P(X = 1)$ y $P(X = 4)$, suponiendo que X tiene una distribución binomial, y compare los resultados con los obtenidos con la distribución hipergeométrica.

c) Para el ejercicio 4-75, calcule $P(X = 1)$ y $P(X = 4)$, suponiendo que X tiene una distribución binomial, y compare los resultados con los obtenidos con la distribución hipergeométrica.

4-86. Utilice la aproximación binomial de la distribución hipergeométrica para aproximar las probabilidades del ejercicio 4-79. ¿Cuál es la corrección para poblaciones finitas en este ejercicio?

4-9 DISTRIBUCIÓN DE POISSON

La distribución de Poisson se introduce con un ejemplo.

EJEMPLO 4-29

Considere la transmisión de n bits en un canal de comunicación digital. Sea la variable aleatoria X igual al número de bits con error. Cuando la probabilidad de que un bit tenga un error es constante y las transmisiones son independientes, X tiene una distribución binomial. Sea que p denote la probabilidad de que un bit tenga un error. Entonces, $E(X) = pn$. Ahora bien, suponga que el número de bits transmitidos aumenta y que la probabilidad de un error disminuye justo lo suficiente para que pn se mantenga igual a una constante, digamos λ. Es decir, n se incrementa y p decrece consecuentemente, de tal modo que $E(X)$ permanece constante. Entonces,

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x}\left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Con cierto trabajo, puede demostrarse que

$$\lim_{n \to \infty} P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 1, 2, \ldots$$

Además, debido a que el número de bits transmitidos tiende a infinito, el número de errores puede ser igual a cualquier entero no negativo. Por lo tanto, el rango de X son los enteros desde cero hasta infinito.

La distribución obtenida como el límite en el ejemplo anterior es más útil de lo que implica la deducción anterior. En el ejemplo siguiente se ilustra la amplitud de sus aplicaciones.

EJEMPLO 4-30

Se presentan imperfecciones aleatoriamente a lo largo de un alambre delgado de cobre. Sea que X denote la variable aleatoria que cuenta el número de imperfecciones en una longitud de L milímetros de alambre, y suponga que el número promedio de imperfecciones en L milímetros es λ.

La distribución de probabilidad de X puede encontrarse con un razonamiento similar al del ejemplo anterior. Se hace la partición de la longitud del alambre en n subintervalos de longitud pequeña, digamos, de 1 micrón (milésima de milímetro) cada uno. Si el subintervalo escogido es lo suficientemente pequeño, la probabilidad de que ocurra más de una imperfección en el subintervalo es despreciable. Además, del supuesto de que las imperfecciones ocurren aleatoriamente, puede interpretarse el que esto implica que cada subintervalo tiene la misma probabilidad de contener una imperfección, digamos, p. Por último, si se supone que la probabilidad de que un subintervalo contenga una imperfección es independiente de otros subintervalos, entonces el modelo de la distribución de X puede establecerse aproximadamente como una variable aleatoria binomial. Puesto que

$$E(X) = \lambda = np$$

se obtiene

$$p = \lambda/n$$

Es decir, la probabilidad de que un subintervalo contenga una imperfección es λ/n. Con subintervalos lo suficientemente pequeños, n es muy grande y p es muy pequeña. Por lo tanto, la distribución de X se obtiene como en el ejemplo anterior.

El ejemplo 4-30 puede generalizarse para incluir un amplio espectro de experimentos aleatorios. El intervalo en el que se hizo la partición, era la longitud de un alambre. Sin embargo, puede aplicarse el mismo razonamiento a cualquier intervalo, incluyendo un intervalo de tiempo, un área o un volumen. Por ejemplo, se ha establecido un modelo de éxito para el conteo de 1) las partículas de contaminación en la fabricación de semiconductores, 2) las imperfecciones en rollos de tela, 3) las llamadas a una central telefónica, 4) las interrupciones de la energía eléctrica y 5) las partículas atómicas emitidas por una muestra, con la función de masa de probabilidad en la definición siguiente:

Definición

Dado un intervalo de números reales, suponga que ocurren conteos al azar a lo largo del intervalo. Si puede hacerse la partición del intervalo en subintervalos con una longitud suficientemente pequeña tal que

1) la probabilidad de más de un conteo en un subintervalo es cero,
2) la probabilidad de un conteo en un subintervalo es la misma para todos los subintervalos y proporcional a la longitud del subintervalo, y
3) el conteo en cada subintervalo es independiente de los demás subintervalos,

entonces el experimento aleatorio se denomina *proceso de Poisson.*

Si el número promedio de conteos en el intervalo es $\lambda > 0$, la variable aleatoria X, que es igual al número de conteos en el intervalo, tiene una **distribución de Poisson** con parámetro λ, y la función de masa de probabilidad de X es

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots \tag{4-13}$$

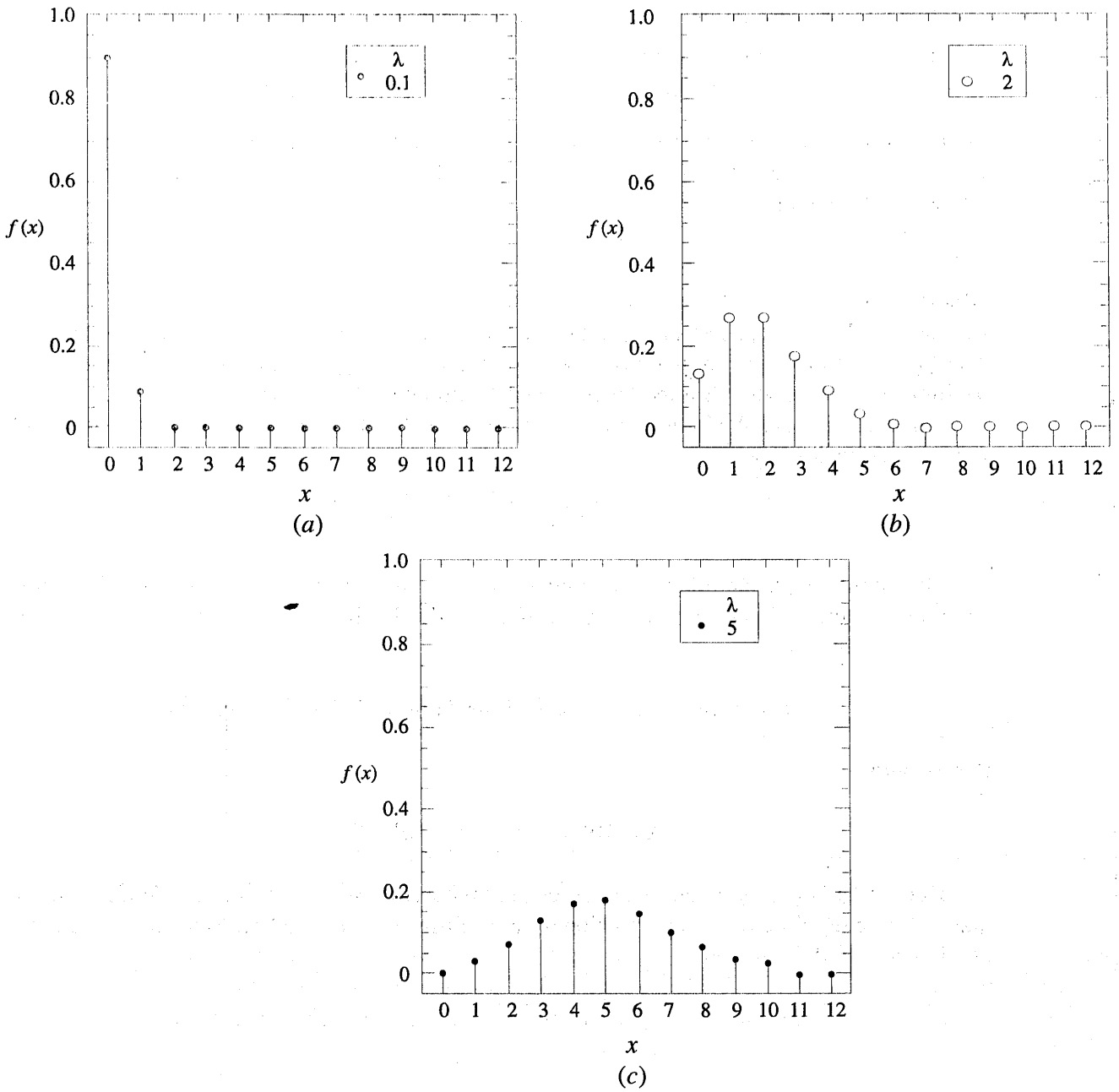


Figura 4-12 Distribución de Poisson para valores seleccionados de los parámetros.

Históricamente, el término *proceso* se ha usado para sugerir la observación de un sistema en el tiempo. En el ejemplo con el alambre de cobre, se demostró que la distribución de Poisson también podía aplicarse a intervalos tales como longitudes. En la figura 4-12 se muestran gráficas de distribuciones de Poisson escogidas.

Es importante **usar unidades consistentes** en el cálculo de probabilidades, medias y varianzas cuando se trabaja con variables aleatorias de Poisson. En el ejemplo siguiente se ilustra la conversión de unidades. Por ejemplo, si el

número promedio de imperfecciones por milímetro de alambre es 3.4, entonces el número promedio de imperfecciones en 10 milímetros de alambre es 34, y el número promedio de imperfecciones en 100 milímetros de alambre es 340.

Si una variable aleatoria de Poisson representa el número de conteos en cierto intervalo, entonces la media de la variable aleatoria debe ser igual al número esperado de conteos en la misma longitud del intervalo.

EJEMPLO 4-31

Para el caso del alambre delgado de cobre, suponga que el número de imperfecciones sigue una distribución de Poisson con una media de 2.3 imperfecciones por milímetro. Determine la probabilidad de exactamente 2 imperfecciones en 1 milímetro de alambre.

Sea que X denote el número de imperfecciones en 1 milímetro de alambre. Entonces, $E(X) = 2.3$ imperfecciones y

$$P(X=2)=\frac{e^{-2.3}2.3^2}{2!}=0.265$$

Determine la probabilidad de 10 imperfecciones en 5 milímetros de alambre. Sea que X denote el número de imperfecciones en 5 milímetros de alambre. Entonces, X tiene una distribución de Poisson con

$$E(X) = 5 \text{ mm} \times 2.3 \text{ imperfecciones/mm} = 11.5 \text{ imperfecciones}$$

Por lo tanto,

$$P(X = 10) = e^{-115}\, 11.5^{10}/10! = 0.113$$

Determine la probabilidad de al menos una imperfección en 2 milímetros de alambre. Sea que X denote el número de imperfecciones en 2 milímetros de alambre. Entonces, X tiene una distribución de Poisson con

$$E(X) = 2 \text{ mm} \times 2.3 \text{ imperfecciones/mm} = 4.6 \text{ imperfecciones}$$

Por lo tanto,

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - e^{-4.6} \\ &= 0.9899 \end{aligned}$$

EJEMPLO 4-32

La contaminación constituye un problema en la fabricación de discos de almacenamiento óptico. El número de partículas de contaminación que ocurre en un disco óptico tiene una distribución de Poisson y el número promedio de partículas por centímetro cuadrado de superficie del disco es 0.1. El área de un disco bajo estudio es 100 centímetros cuadrados. Encuentre la probabilidad de que ocurran 12 partículas en el área del disco bajo estudio.

Sea que X denote el número de partículas en el área de un disco bajo estudio. Puesto que el número promedio de partículas es 0.1 partículas por cm^2

$$E(X) = 100\ cm^2 \times 0.1\ \text{partículas/cm}^2$$
$$= 10\ \text{partículas}$$

Por lo tanto,

$$P(X = 12) = \frac{e^{-10}10^{12}}{12!}$$
$$= 0.095$$

La probabilidad de ocurran cero partículas en el área del disco bajo estudio es

$$P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$$

Determine la probabilidad de que 12 o menos partículas ocurran en el área del disco bajo estudio. La probabilidad es

$$P(X \leq 12) = P(X = 0) + P(X = 1) + \ldots + P(X = 12) = \sum_{i=0}^{12} \frac{e^{-10}10^{i}}{i!}$$

Debido a que el cálculo de esta suma es muy laborioso, muchos programas de computadora calculan probabilidades de Poisson acumuladas. Por uno de tales programas, $P(X \leq 12) = 0.791$.

La deducción de la media y la varianza de una variable aleatoria de Poisson se deja como ejercicio. Los resultados son los siguientes:

Si X es una variable aleatoria de Poisson con parámetro λ, entonces la media y la varianza de X son

$$\mu = E(X) = \lambda \quad \text{y} \quad \sigma^2 = V(X) = \lambda \tag{4-14}$$

Por ejemplo, si los conteos de partículas siguen una distribución de Poisson con una media de 25 partículas por centímetro cuadrado, entonces la desviación estándar de los conteos es 5 por centímetro cuadrado. Por consiguiente, la información sobre la variabilidad se obtiene con facilidad. A la inversa, si la varianza de los datos de los conteos es mucho más grande que la media de los mismos datos, entonces la distribución de Poisson no es un buen modelo para la distribución de la variable aleatoria.

EJERCICIOS DE LA SECCIÓN 4-9

4-84. Suponga que X tiene una distribución de Poisson con una media de 4. Determine las siguientes probabilidades:

a) $P(X = 0)$
b) $P(X \leq 2)$
c) $P(X = 4)$
d) $P(X = 8)$

4-85. Suponga que X tiene una distribución de Poisson con una media de 0.4. Determine las siguientes probabilidades.

a) $P(X = 0)$
b) $P(X \leq 2)$
c) $P(X = 4)$
d) $P(X = 8)$

4-86. Suponga que el número de clientes que entran a un banco en una hora es una variable aleatoria de Poisson y suponga que $P(X = 0) = 0.05$. Determine la media y la varianza de X.

4-87. Con frecuencia se establece un modelo de Poisson para el número de llamadas telefónicas que entran en una central telefónica. Suponga que hay en promedio 10 llamadas por hora.

a) ¿Cuál es la probabilidad de que haya exactamente 5 llamadas en una hora?
b) ¿Cuál es la probabilidad de que haya 3 o menos llamadas en una hora?
c) ¿Cuál es la probabilidad de que haya exactamente 15 llamadas en dos horas?
d) ¿Cuál es la probabilidad de que haya exactamente 5 llamadas en 30 minutos?

4-88. Se supone que el número de imperfecciones en los rollos de tela de una fábrica textil tiene una distribución de Poisson con una media de 0.1 imperfecciones por metro cuadrado.

a) ¿Cuál es la probabilidad de que haya dos imperfecciones en 1 metro cuadrado de tela?
b) ¿Cuál es la probabilidad de que haya una imperfección en 10 metros cuadrados de tela?
c) ¿Cuál es la probabilidad de que no haya imperfecciones en 20 metros cuadrados de tela?
d) ¿Cuál es la probabilidad de que haya al menos dos imperfecciones en 10 metros cuadrados de tela?

4-89. Se supone que el número de grietas en un tramo de una carretera interestatal que tienen la suficiente importancia para requerir reparación sigue una distribución de Poisson con una media de dos grietas por milla.

a) ¿Cuál es la probabilidad de que no haya grietas que requieran reparación en 5 millas de carretera?
b) ¿Cuál es la probabilidad de que haya al menos una grieta que requiera reparación en media milla de carretera?
c) Si el número de grietas se relaciona con la carga vehicular en la carretera y algunos tramos de la misma tienen una carga intensa de vehículos en tanto que en otros tramos la carga vehicular es ligera, ¿qué piensa el lector acerca del supuesto de que el número de grietas que requieren reparación sigue una distribución de Poisson?

4-90. El número de imperfecciones superficiales en los tableros de plástico utilizados en el interior de automóviles tiene una distribución de Poisson con una media de 0.05 imperfecciones por pie cuadrado de tablero de plástico. Suponga que el interior de un automóvil contiene 10 pies cuadrados de tablero de plástico.

a) ¿Cuál es la probabilidad de que no haya imperfecciones superficiales en el interior de un automóvil?
b) Si se venden 10 unidades a una compañía de renta, ¿cuál es la probabilidad de que ninguno de los 10 autos tenga alguna imperfección superficial?
c) Si se venden 10 autos a una compañía de renta, ¿cuál es la probabilidad de que a lo sumo uno de los autos tenga alguna imperfección superficial?

4-91. El número de fallas de un instrumento de prueba de partículas de contaminación en un producto es una variable aleatoria de Poisson con una media de 0.02 fallas por hora.

a) ¿Cuál es la probabilidad de que el instrumento no falle en una corrida de 8 horas?
b) ¿Cuál es la probabilidad de al menos una falla en un día de 24 horas?

Ejercicios complementarios

4-92. Un embarque de sustancias químicas llega en 15 contenedores. Tres de los contenedores se seleccionan, al azar y sin reemplazo, para una inspección de pureza. Si dos de los contenedores no cumplen con los requerimientos de pureza, ¿cuál es la probabilidad de que al menos uno de los contenedores que no cumplen se seleccione en la muestra?

4-93. La probabilidad de que una llamada a una línea de servicios sea contestada en menos de 30 segundos es 0.75. Suponga que las llamadas son independientes.
a) Si se hacen 10 llamadas, ¿cuál es la probabilidad de que exactamente 9 se contesten dentro de 30 segundos?
b) Si se hacen 20 llamadas, ¿cuál es la probabilidad de que al menos 16 se contesten en menos de 30 segundos?
c) Si se hacen 20 llamadas, ¿cuál es el número promedio de llamadas que se contestan en menos de 30 segundos?

4-94. Continuación del ejercicio 4-93.
a) ¿Cuál es la probabilidad de que tengan que hacerse cuatro llamadas para obtener la primera contestación en menos de 30 segundos?
b) ¿Cuál es el número promedio de llamadas hasta obtener una contestación en menos de 30 segundos?

4-95. Continuación del ejercicio 4-94.
a) ¿Cuál es la probabilidad de que tengan que hacerse seis llamadas para que dos se contesten en menos de 30 segundos?
b) ¿Cuál es el número promedio de llamadas para obtener dos contestaciones en menos de 30 segundos?

4-96. El número de mensajes enviados a un buzón electrónico es una variable aleatoria de Poisson con una media de 5 mensajes por hora.
a) ¿Cuál es la probabilidad de que se reciban 5 mensajes en una hora?
b) ¿Cuál es la probabilidad de que se reciban 10 mensajes en 1.5 horas?
c) ¿Cuál es la probabilidad de que se reciban menos de dos mensajes en media hora?

4-97. El número de erratas en un libro de texto sigue una distribución de Poisson con una media de 0.01 erratas por página. ¿Cuál es la probabilidad de que haya tres erratas o menos en 100 páginas?

4-98. La probabilidad de que un individuo se recupere de una enfermedad en el transcurso de una semana sin tratamiento es 0.1. Suponga que 20 individuos independientes que padecen esta enfermedad reciben tratamiento con un medicamento y 4 se recuperan en el transcurso de una semana. Si el medicamento no tiene ningún efecto, ¿cuál es la probabilidad de que 4 personas o más se recuperen en el transcurso de una semana?

4-99. En un proceso de manufactura en que se laminan varias capas de cerámica, 1% de las unidades ensambladas están defectuosas. Suponga que las unidades ensambladas son independientes.
a) ¿Cuál es el número promedio de unidades ensambladas que es necesario verificar para obtener cinco unidades ensambladas defectuosas?
b) ¿Cuál es la desviación estándar del número de unidades ensambladas que es necesario verificar para obtener cinco unidades ensambladas defectuosas?

4-100. Continuación del ejercicio 4-99. Determinar el número mínimo de unidades ensambladas que es necesario verificar para que la probabilidad de al menos una unidad ensamblada defectuosa exceda 0.95.

4-101. Determine la constante c para que la función siguiente sea una función de masa de probabilidad: $f(x) = cx$ para $x = 1, 2, 3, 4$.

4-102. Un fabricante de productos electrónicos de consumo espera que 2% de las unidades fallen durante el periodo de garantía. Se hace

el seguimiento del cumplimiento de la garantía en una muestra de 500 unidades independientes.

a) ¿Cuál es la probabilidad de que ninguna unidad falle durante el periodo de garantía?

b) ¿Cuál es el número esperado de fallas durante el periodo de garantía?

c) ¿Cuál es la probabilidad de que más de dos unidades fallen durante el periodo de garantía?

4-103. La variable aleatoria X tiene la siguiente distribución de probabilidad.

x	2	3	5	8
probabilidad	0.2	0.4	0.3	0.1

Determine lo siguiente:

a) $P(X \leq 3)$

b) $P(X > 2.5)$

c) $P(2.7 < X < 5.1)$

d) $E(X)$

e) $V(X)$

4-104. Determine la función de masa de probabilidad para la variable aleatoria con la siguiente función de distribución acumulada.

$$F(x) = \begin{cases} 0 & x < 2 \\ 0.2 & 2 \leq x < 5.7 \\ 0.5 & 5.7 \leq x < 6.5 \\ 0.8 & 6.5 \leq x < 8.5 \\ 1 & 8.5 \leq x \end{cases}$$

4-105. La tapa de cada cojinete principal de un motor tiene cuatro tornillos. Los tornillos se seleccionan, al azar y sin reemplazo, de un cajón de partes de repuesto que contiene 30 tornillos de un proveedor y 70 tornillos de otro.

a) ¿Cuál es la probabilidad de que en una tapa del cojinete principal, todos los tornillos sean del mismo proveedor?

b) ¿Cuál es la probabilidad de que exactamente tres tornillos sean del mismo proveedor?

4-106. Suponga que el número de errores en una superficie de grabación magnética es una variable aleatoria de Poisson con una media de un error cada 10^5 bits. Un sector de datos consta de 4 096 bytes de ocho bits.

a) ¿Cuál es la probabilidad de más de un error en un sector?

b) ¿Cuál es el número promedio de sectores hasta encontrar un error?

4-107. Un técnico de instalación para un sistema especializado de comunicaciones se envía a una ciudad únicamente cuando se han hecho tres o más pedidos. Suponga que los pedidos siguen una distribución de Poisson con una media de 0.25 por semana para una ciudad con 100 000 habitantes y suponga que la ciudad donde vive el lector tiene 800 000 habitantes.

a) ¿Cuál es la probabilidad de que se necesite un técnico después de un periodo de una semana?

b) Si usted es la primera persona de la ciudad que hace un pedido, ¿cuál es la probabilidad de que tenga que esperar más de dos semanas a partir del momento en que hizo su pedido hasta que se envíe un técnico?

4-108. Se sospecha que algunas de las cargas que contienen sustancias químicas compradas a un proveedor exceden el contenido de humedad permitido. Se van a probar muestras de 30 cargas para el contenido de humedad. Suponga que las cargas son independientes. Determine la proporción de cargas del proveedor que deben exceder el contenido de humedad permitido, para que sea 0.90 la probabilidad de que al menos un contenedor de la muestra de 30 no pase la prueba.

4-109. Llegan mensajes a un servidor de computadora de acuerdo con una distribución de Poisson con una rapidez media de 10 por hora. Determine la longitud del intervalo de tiempo para que la probabilidad de que no lleguen mensajes durante este intervalo sea 0.90.

4-110. Se presentan imperfecciones en los tableros de plástico usados en el interior de automóviles de acuerdo con una distribución de Poisson con una media de 0.02 imperfecciones por tablero.

a) Si se inspeccionan 50 tableros, ¿cuál es la probabilidad de que no haya imperfecciones?

b) ¿Cuál es el número esperado de tableros que es necesario inspeccionar antes de encontrar una imperfección?

c) Si se inspeccionan 50 tableros, ¿cuál es la probabilidad de que el número de tableros que tienen dos o más imperfecciones sea menor o igual que dos?

EJERCICIOS PARA DESARROLLAR EL INTELECTO

4-111. Deduzca la media y la varianza de una variable aleatoria hipergeométrica (ejercicio difícil).

4-112. Deduzca la fórmula para la media y la desviación estándar de una variable aleatoria discreta uniforme en el rango de los enteros $a, a + 1, ..., b$.

4-113. Una compañía inspecciona los envíos de los proveedores a fin de detectar productos que no cumplen con las especificaciones. Suponga que un lote contiene 1 000 artículos y que 1% no cumple. ¿Cuál es el tamaño de la muestra que se necesita para que la probabilidad de escoger al menos un artículo de la muestra que no cumple, sea al menos 0.90? Suponga que es adecuada la aproximación binomial de la distribución hipergeométrica.

4-114. Una compañía inspecciona los envíos de los proveedores a fin de detectar productos que no cumplen con las especificaciones. La política de la compañía es utilizar siempre un tamaño de la muestra que es 10% del tamaño del lote. Comente acerca de la efectividad de esta política como regla general para todos los tamaños de los lotes.

4-115. Las imperfecciones superficiales en los tableros exteriores de automóviles siguen una distribución de Poisson con una media de 0.1 imperfecciones por tablero. Si se verifican 100 tableros, ¿cuál es la probabilidad de que menos de 5 tableros tengan alguna imperfección?

4-116. Una panificadora grande puede producir panecillos en lotes de 0, 1 000, 2 000 o 3 000 por día. El costo de producción por pieza es \$0.10. La demanda varía aleatoriamente de acuerdo con la distribución siguiente:

demanda de panecillos	0	1 000	2 000	3 000
probabilidad de la demanda	0.3	0.2	0.3	0.2

Cada panecillo para el que hay demanda se vende en \$0.30. Cada panecillo para el que no hay demanda se vende en un mercado de segunda en \$0.05. ¿Cuántos panecillos deberá producir la panificadora diariamente para maximizar las ganancias promedio?

4-117. Un fabricante almacena componentes obtenidos de un proveedor. Suponga que 2% de los componentes están defectuosos y que los componentes defectuosos ocurren independientemente. ¿Cuántos componentes debe tener almacenados el fabricante para que la probabilidad de que puedan surtirse 100 pedidos, sin reordenar componentes, sea al menos 0.95?

5 Variables aleatorias continuas y distribuciones de probabilidad

PLAN GENERAL DEL CAPÍTULO

5-1 VARIABLES ALEATORIAS CONTINUAS

5-2 DISTRIBUCIONES DE PROBABILIDAD Y FUNCIONES DE DENSIDAD DE PROBABILIDAD

5-3 FUNCIONES DE DISTRIBUCIÓN ACUMULADA

5-4 MEDIA Y VARIANZA DE UNA VARIABLE ALEATORIA CONTINUA

5-5 DISTRIBUCIÓN CONTINUA UNIFORME

5-6 DISTRIBUCIÓN NORMAL

5-7 GRÁFICAS DE PROBABILIDAD

5-8 APROXIMACIÓN NORMAL A LA DISTRIBUCIÓN BINOMIAL Y DE POISSON

5-9 DISTRIBUCIÓN EXPONENCIAL

5-10 DISTRIBUCIONES DE ERLANG Y GAMMA

5-10.1 Distribución de Erlang

5-10.2 Distribución gamma

5-11 DISTRIBUCIÓN DE WEIBULL

5-1 VARIABLES ALEATORIAS CONTINUAS

En el capítulo anterior se comentó la medición de la corriente en un alambre delgado de cobre. Se señaló que los resultados podrían diferir ligeramente en mediciones hechas en días diferentes debido a las pequeñas variaciones en las variables del experimento que no están controladas (cambios en la temperatura ambiente, pequeñas impurezas en la composición química del alambre, desviaciones en la fuente de energía, etcétera).

Otro ejemplo es, en un día de producción, seleccionar una pieza y medir con precisión una longitud dimensional. En la práctica puede haber ligeras variaciones en las longitudes reales

medidas debido a diversas causas, como vibraciones, fluctuaciones en la temperatura, operadores diferentes, calibración, desgaste de la herramienta de corte, uso de los asientos o cambio de materias primas. Incluso el procedimiento de medición puede dar lugar a variaciones en los resultados finales.

En este tipo de experimentos, la medición de interés —la corriente en el experimento con un alambre de cobre, la longitud de una pieza maquinada— puede representarse por una variable aleatoria. Es razonable modelar el rango de los valores posibles de la variable aleatoria mediante un intervalo (finito o infinito) de números reales. Por ejemplo, para la longitud de la pieza maquinada, el modelo permite que la medición del experimento tenga como resultado cualquier valor en un intervalo de números reales. Debido a que el rango es cualquier valor de un intervalo, el modelo permite cualquier grado de precisión en mediciones de longitud. Sin embargo, debido a que el número de valores posibles de la variable aleatoria X es finito incontable, X posee una distribución claramente diferente de las variables aleatorias discretas estudiadas con anterioridad. El rango de X incluye todos los valores en un intervalo de números reales; es decir, el rango de X puede considerarse un continuo.

5-2 DISTRIBUCIONES DE PROBABILIDAD Y FUNCIONES DE DENSIDAD DE PROBABILIDAD

En la ingeniería es común el uso de las funciones de densidad para describir sistemas físicos. Por ejemplo, considérese la densidad de una distribución de cargas sobre una viga larga y delgada, como se muestra en la figura 5-1. Para cualquier punto x sobre la viga, la densidad puede describirse con una función (en gramos/cm). Los intervalos con distribución de cargas grandes corresponden a valores grandes de la función. La distribución de carga total entre los puntos a y b está determinada por la integral de la función de densidad de a a b. Esta integral es el área bajo la función de densidad sobre el intervalo y en un sentido amplio puede interpretarse como la suma de todas las cargas en este intervalo.

De manera similar, puede usarse una **función de densidad de probabilidad** $f(x)$ para describir la distribución de probabilidad de una variable aleatoria continua X. Si es factible que un intervalo contenga un valor de X, entonces su probabilidad es grande y corresponde a valores grandes de $f(x)$. La probabilidad de que X esté entre a y b se determina como la integral de $f(x)$ de a a b (véase la figura 5-2).

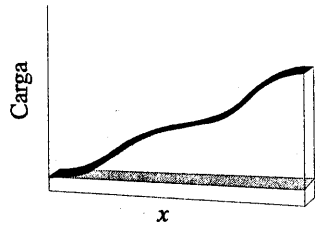


Figura 5-1 Función de densidad de una distribución de carga sobre una viga larga y delgada.

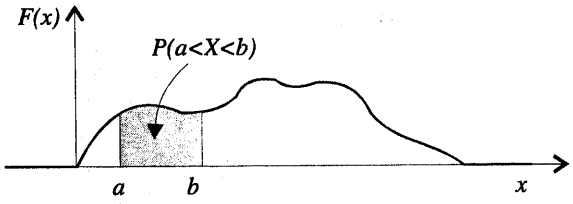


Figura 5-2 Probabilidad determinada por el área bajo $f(x)$.

Definición

Para una variable aleatoria continua X, una **función de densidad de probabilidad** es una función tal que

1) $f(x) \geq 0$

2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

3) $P(a \leq X \leq b) = \int_a^b f(x)\,dx =$ área bajo $f(x)$ de a a b, para a y b cualesquiera (5-1)

Una función de densidad de probabilidad proporciona una descripción simple de las probabilidades asociadas con una variable aleatoria. En tanto $f(x)$ sea no negativa y $\int_{-\infty}^{\infty} f(x)\,dx = 1$, $0 \leq P(a < X < b) \leq 1$, por lo que las probabilidades se encuentran restringidas de manera conveniente. Una función de densidad de probabilidad es cero para los valores de x que no pueden ocurrir y se supone que es cero siempre que no se defina expresamente.

Un histograma es una aproximación de una función de densidad de probabilidad (véase la figura 5-3). Para cada intervalo del histograma, el área del rectángulo es igual a la frecuencia relativa (proporción) de las mediciones en el intervalo. La frecuencia relativa es una estimación de la probabilidad de que una medición esté en el intervalo. De manera similar, el área bajo $f(x)$ en cualquier intervalo es igual a la probabilidad real de que una medición esté en el intervalo.

El punto importante es que **$f(x)$ se usa para calcular un área** que representa la probabilidad de que X asuma un valor en $[a, b]$. Para el ejemplo de la medición de la corriente eléctrica, la probabilidad de que X esté en el intervalo [14 mA, 15 mA] es la integral de la función de densidad de probabilidad de X, $f(x)$, en este intervalo. La probabilidad de que X esté en el intervalo [14.5 mA, 14.6 mA] es la integral de la misma función, $f(x)$, en este intervalo de menor longitud. Mediante la elección apropiada de la forma de $f(x)$, es posible representar las probabilidades asociadas con cualquier variable aleatoria X. La forma de $f(x)$ determina la manera en que se compara la probabilidad de que X asuma un valor en el intervalo [14.5 mA, 14.6 mA] con la probabilidad de cualquier otro intervalo de longitud igual o diferente.

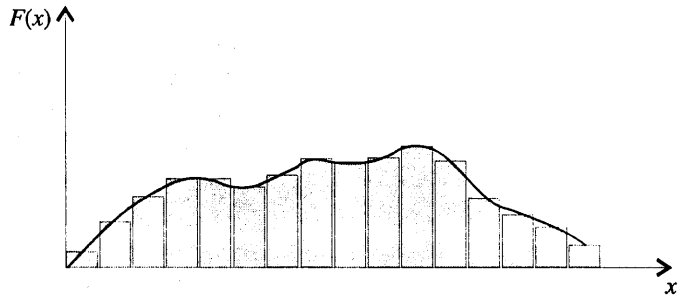


Figura 5-3 Un histograma es una aproximación de una función de densidad de probabilidad.

Para la función de densidad de una distribución de carga sobre una viga larga y delgada, debido a que cualquier punto tiene longitud cero, la carga en cualquier punto es cero. De manera similar, para una variable aleatoria continua X y *cualquier* valor x,

$$P(X = x) = 0$$

Con base en este resultado, podría parecer que nuestro modelo de una variable aleatoria continua es inútil. Sin embargo, en la práctica, cuando se hace una medición particular de la corriente, como 14.47 miliamperes, este resultado puede interpretarse como el valor redondeado de una medición de la corriente que está en realidad en el rango $14.465 \leq x \leq 14.475$. Por lo tanto, la probabilidad de que se observe el valor redondeado 14.47 como valor de X es la probabilidad de que X asuma un valor en el intervalo [14.465, 14.475], que es diferente de cero. De manera similar, nuestro modelo para una variable aleatoria continua implica lo siguiente:

Si X es una variable aleatoria continua, entonces para cualesquiera x_1 y x_2,

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2) \qquad (5\text{-}2)$$

EJEMPLO 5-1

Sea que la variable aleatoria continua X denote la corriente medida en un alambre delgado de cobre en miliamperes. Suponga que el rango de X es [0, 20 mA], y suponga asimismo que la función de densidad de probabilidad de X es $f(x) = 0.05$ para $0 \leq x \leq 20$. ¿Cuál es la probabilidad de que una medición de la corriente sea menor que 10 miliamperes?

En la figura 5-4 se muestra la función de densidad de probabilidad. Se supone que $f(x) = 0$ siempre que no está definida expresamente. La probabilidad pedida se indica con el área sombreada en la figura 5-4.

$$P(X < 10) = \int_0^{10} f(x)\,dx = 0.5$$

Como otro ejemplo,

$$P(5 < X < 15) = \int_5^{15} f(x)\,dx = 0.5$$

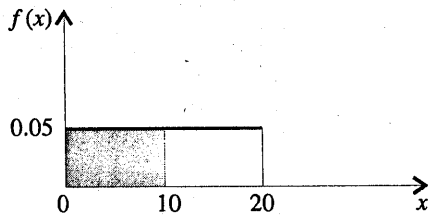


Figura 5-4 Función de densidad de probabilidad del ejemplo 5-1.

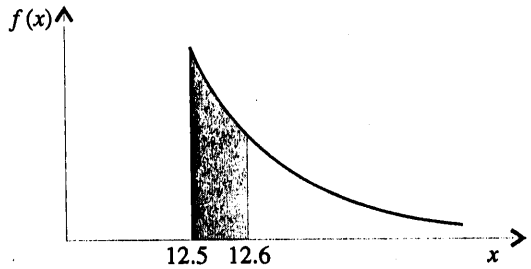


Figura 5-5 Función de densidad de probabilidad del ejemplo 5-2.

EJEMPLO 5-2

Sea que la variable aleatoria continua X denote el diámetro de un agujero taladrado en un componente metálico. El diámetro especificado es 12.5 milímetros. La mayoría de las perturbaciones aleatorias del proceso resultan en diámetros mayores. Datos históricos indican que la distribución de X puede modelarse con la función de densidad de probabilidad $f(x) = 20e^{-20(x-12.5)}$, $x \geq 12.5$.

Si un componente con un diámetro mayor que 12.60 milímetros se desecha, ¿qué proporción de los componentes se desecha? La función de densidad y la probabilidad pedida se muestran en la figura 5-5. Un componente se desecha si $X > 12.60$. Entonces,

$$P(X>12.60)=\int_{12.6}^{\infty} f(x)dx=\int_{12.6}^{\infty} 20e^{-20(x-12.5)}dx=-e^{-20(x-12.5)}\Big|_{12.6}^{\infty}=0.135$$

¿Qué proporción de los componentes tiene entre 12.5 y 12.6 milímetros? Entonces,

$$P(12.5<X<12.6)=\int_{12.5}^{12.6} f(x)dx=-e^{-20(x-12.5)}\Big|_{12.5}^{12.6}=0.865$$

Como el área total bajo $f(x)$ es igual a 1, también puede calcularse $P(12.5 < X < 12.6) = 1 - P(X > 12.6) = 1 - 0.135 = 0.865$.

EJERCICIOS DE LA SECCIÓN 5-2

5-1. Suponga que $f(x) = e^{-x}$ para $0 < x$. Determine las siguientes probabilidades:
a) $P(1 < X)$
b) $P(1 < X < 2.5)$
c) $P(X = 3)$
d) $P(X < 4)$
e) $P(3 \leq X)$

5-2. Suponga que $f(x) = e^{-x}$ para $0 < x$.
a) Determine x tal que $P(x < X) = 0.10$.
b) Determine x tal que $P(X \leq x) = 0.10$.

5-3. Suponga que $f(x) = e^{-(x-4)}$ para $4 < x$. Determine las siguientes probabilidades:
a) $P(1 < X)$
b) $P(2 \leq X < 5)$
c) $P(5 < X)$
d) $P(8 < X < 12)$
e) Determine x tal que $P(X < x) = 0.90$.

5-4. Suponga que $f(x) = 1.5x^2$ para $-1 < x < 1$. Determine las siguientes probabilidades:
a) $P(0 < X)$

b) $P(0.5 < X)$
c) $P(-0.5 \leq X \leq 0.5)$
d) $P(X < -2)$
e) $P(X < 0 \text{ o } X > -0.5)$
f) Determine x tal que $P(x < X) = 0.05$.

5-5. La función de densidad de probabilidad del del tiempo de falla de un componente electrónico en una copiadora (en horas) es $f(x) = \dfrac{e^{-x/1\,000}}{1\,000}$ para $x > 0$. Determine la probabilidad de que

a) un componente dure más de 3 000 horas antes de fallar,
b) un componente falle en el intervalo de 1 000 a 2 000 horas y
c) un componente falle antes de 1 000 horas.
d) Determine el número de horas en que 10% de los componentes han fallado.

5-6. La función de densidad de probabilidad del peso neto en libras de un herbicida químico empacado es $f(x) = 2.0$ para $49.75 < x < 50.25$ libras.

a) Determine la probabilidad de que un paquete pese más de 50 libras.
b) ¿Cuánto herbicida contiene el 90% de los paquetes?

5-7. La función de densidad de probabilidad de la longitud de una bisagra para sostener puertas es $f(x) = 1.25$ para $74.6 < x < 75.4$ milímetros. Determine lo siguiente:

a) $P(X < 74.8)$
b) $P(X < 74.8 \text{ o } X > 75.2)$
c) Si las especificaciones para este proceso son de 74.7 a 75.3 milímetros, ¿qué proporción de las bisagras cumple con las especificaciones?

5-8. La función de densidad de probabilidad de la longitud de una barra de metal es $f(x) = 2$ para $2.3 < x < 2.8$ metros.

a) Si las especificaciones para este proceso son de 2.25 a 2.75 metros, ¿qué proporción de las barras no cumple con las especificaciones?
b) Suponga que la función de densidad de probabilidad para un intervalo de longitud de 0.5 metros, es $f(x) = 2$. ¿Cuál deberá ser el valor central de la densidad para obtener la proporción mayor de varillas, dentro de las especificaciones?

5-9. Si X es una variable aleatoria continua, demuestre que $P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2)$.

5-3 FUNCIONES DE DISTRIBUCIÓN ACUMULADA

Un método alternativo para describir la distribución de una variable aleatoria discreta también puede usarse para variables aleatorias continuas.

Definición

La **función de distribución acumulada** de una variable aleatoria continua X es

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \qquad (5\text{-}3)$$

para $-\infty < x < \infty$.

Amplíe la definición de $f(x)$ a la recta real completa permite definir la función de distribución acumulada para todos los números reales. En el ejemplo siguiente se ilustra la definición.

EJEMPLO 5-3

Para la medición de la corriente en el alambre de cobre del ejemplo 5-1, la función de distribución acumulada de la variable aleatoria X consta de tres expresiones. Si $x < 0$, entonces $f(x) = 0$. Por lo tanto,

$$F(x) = 0, \text{ para } x < 0$$

y

$$F(x) = \int_0^x f(u)du = 0.05x, \text{ para } 0 \le x < 20$$

Por último,

$$F(x) = \int_0^x f(u)du = 1, \text{ para } 20 \le x$$

Por lo tanto,

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.05x & 0 \le x < 20 \\ 1 & 20 \le x \end{cases}$$

En la figura 5-6 se muestra la gráfica de $F(x)$.

Obsérvese que en la definición de $F(x)$ cualquier signo $<$ puede cambiarse por $\le$ y viceversa. Es decir, $F(x)$ puede definirse como $0.05x$ o bien como 0 en el punto final $x = 0$, y $F(x)$ puede definirse como $0.05x$ o bien como 1 en el punto final $x = 20$. En otras palabras, $F(x)$ es una función continua. Para una variable aleatoria discreta, $F(x)$ no es una función continua. En ocasiones, una variable aleatoria continua se define como aquella que tiene una función de distribución acumulada continua.

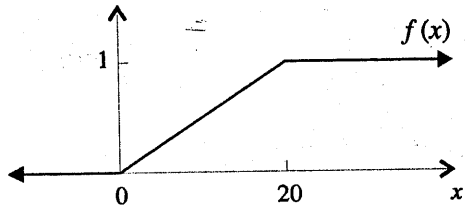


Figura 5-6 Función de distribución acumulada del ejemplo 5-3.

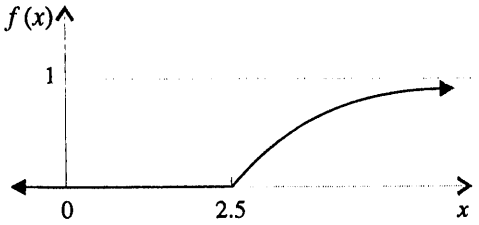


Figura 5-7 Función de distribución acumulada del ejemplo 5-4.

EJEMPLO 5-4

Para la operación de taladrado del ejemplo 5-2, $F(x)$ consta de dos expresiones:

$$F(x) = 0 \text{ para } x < 12.5$$

y

$$F(x) = \int_{12.5}^{x} 20e^{-20(u-12.5)}du$$
$$= 1 - e^{-20(x-12.5)}$$

para $12.5 \leq x$. Por lo tanto,

$$F(x) = \begin{cases} 0 & x < 12.5 \\ 1 - e^{-20(x-12.5)} & 12.5 \leq x \end{cases}$$

En la figura 5-7 se muestra la gráfica de $F(x)$.

La función de densidad de probabilidad de una variable aleatoria continua puede determinarse a partir de la función de distribución acumulada por derivación. Es decir, dada $F(x)$,

$$f(x) = \frac{dF(x)}{dx}$$

en tanto que la derivada exista.

EJEMPLO 5-5

El tiempo para que una reacción química se complete (en milisegundos) se aproxima con la función de distribución acumulada

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-0.01x} & 0 \leq x \end{cases}$$

Determine la función de densidad de probabilidad de X. ¿Qué proporción de las reacciones se completa dentro de 200 milisegundos? Usando el resultado de que la función de densidad de probabilidad es la derivada de $F(x)$, se obtiene

$$f(x)=\begin{cases}0 & x<0\\ 1-e^{-0.01x} & x\le x\end{cases}$$

La probabilidad de que una reacción se complete dentro de 200 milisegundos es

$$P(X<200)=F(200)=1-e^{-2}=0.8647.$$

EJERCICIOS DE LA SECCIÓN 5-3

5-10. Suponga que la función de distribución acumulada de la variable aleatoria X es

$$F(x)=\begin{cases}0 & x<0\\ 0.2x & 0\le x<5\\ 1 & 5\le x\end{cases}$$

Determine lo siguiente:

a) $P(X<2.8)$
b) $P(X>1.5)$
c) $P(X<-2)$
d) $P(X>6)$

5-11. Suponga que la función de distribución acumulada de la variable aleatoria X es

$$F(x)=\begin{cases}0 & x<-2\\ 0.25x+0.5 & -2\le x<2\\ 1 & 2\le x\end{cases}$$

Determine lo siguiente:

a) $P(X<1.8)$
b) $P(X>-1.5)$
c) $P(X<-2)$
d) $P(-1<X<1)$

5-12. Determine la función de distribución acumulada para la distribución del ejercicio 5-1.

5-13. Determine la función de distribución acumulada para la distribución del ejercicio 5-3.

5-14. Determine la función de distribución acumulada para la distribución del ejercicio 5-5. Use la función de distribución acumulada para determinar la probabilidad de que un componente dure más de 3 000 horas sin fallar.

5-15. Determine la función de distribución acumulada para la distribución del ejercicio 5-7. Use la función de distribución acumulada para determinar la probabilidad de que una longitud exceda 75 milímetros.

Determine la función de densidad de probabilidad para cada una de las funciones de distribución acumulada siguientes:

5-16. $F(x)=1-e^{-2x} \qquad x>0$

5-17.

$$F(x)=\begin{cases}0 & x<0\\ 0.2x & 0\le x<4\\ 0.04x+0.64 & 4\le x<9\\ 1 & 9\le x\end{cases}$$

5-18.

$$F(x)=\begin{cases}0 & x<-2\\ 0.25x+0.5 & -2\le x<1\\ 0.5x+0.25 & 1\le x<1.5\\ 1 & 1.5\le x\end{cases}$$

5-19. La anchura del entrehierro es una propiedad importante en una cabeza de grabación magnética. En unidades codificadas, si la anchura es una variable aleatoria continua en el rango de $0<x<2$ con $f(x)=0.5x$, determine la función de distribución acumulada de la anchura del entrehierro.

5-4 MEDIA Y VARIANZA DE UNA VARIABLE ALEATORIA CONTINUA

La media y la varianza de una variable aleatoria continua se define en forma similar a una variable aleatoria discreta. La integración reemplaza a la suma en las definiciones.

Definición

Suponga que X es una variable aleatoria continua con función de densidad de probabilidad $f(x)$.

La **media** o **valor esperado** de X, denotada como μ o $E(X)$, es

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx \tag{5-4}$$

La **varianza** de X, denotada como $V(X)$ o σ^2, es

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2 \tag{5-5}$$

La **desviación estándar** de X es $\sigma = [V(X)]^{1/2}$.

EJEMPLO 5-6

Para la medición de la corriente en el alambre de cobre del ejemplo 5-1, la media de X es

$$E(X) = \int_0^{20} x f(x)\, dx = 0.05x^2/2 \Big|_0^{20} = 10$$

La varianza de X es

$$V(X) = \int_0^{20} (x-10)^2 f(x)\, dx = 0.05(x-10)^3/3 \Big|_0^{20} = 33.33$$

EJEMPLO 5-7

Para la operación de taladrado del ejemplo 5-2, la media de X es

$$E(X) = \int_{12.5}^{\infty} xf(x)\,dx$$

Puede aplicarse la integración por partes para demostrar que

$$E(X) = -xe^{-20(x-12.5)} - \frac{e^{-20(x-12.5)}}{20}\Bigg|_{12.5}^{\infty} = 12.5 + 0.05 = 12.55$$

La varianza de X es

$$V(X) = \int_{12.5}^{\infty} (x-12.55)^2 f(x)\,dx$$

Puede integrarse por partes dos veces para demostrar que

$$V(X) = 0.0025$$

EJERCICIOS DE LA SECCIÓN 5-4

5-20. Suponga que $f(x) = 0.25$ para $0 < x < 4$. Determine la media y la varianza de X.

5-21. Suponga que $f(x) = 0.125x$ para $0 < x < 4$. Determine la media y la varianza de X.

5-22. Suponga que $f(x) = 1.5x^2$ para $-1 < x < 1$. Determine la media y la varianza de X.

5-23. Determine la media y la varianza del peso de los paquetes del ejercicio 5-6.

5-24. El espesor de un recubrimiento conductor en micrones (milésimas de milímetro), tiene una función de densidad de $600x^{-2}$ para $100\ \mu m < x < 120\ \mu m$.

a) Determine la media y la varianza del espesor del recubrimiento.

b) Si el costo del recubrimiento es de \$0.50 por micrón (milésima de milímetro) de espesor, en cada pieza, ¿cuál es el costo promedio del recubrimiento por pieza?

5-25. Suponga que el tamaño de las partículas de contaminación en micrones (milésimas de milímetro) puede modelarse como $f(x) = 2x^{-3}$ para $1 < x$. Determine la media de X.

5-26. Es necesario integrar por partes. La función de densidad de probabilidad del diámetro, en milímetros, de un agujero taladrado es $10e^{-10(x-5)}$ para $x > 5$ mm. Aun cuando el diámetro especificado es de 5 milímetros, las vibraciones, el desgaste de la herramienta y otros perjuicios producen diámetros mayores de 5 milímetros.

a) Determine la media y la varianza del diámetro de los agujeros.

b) Determine la probabilidad de que un diámetro exceda 5.1 milímetros.

5-27. Suponga que la función de densidad de probabilidad de la longitud de los cables de una computadora es $f(x) = 0.1$ de 1 200 a 1 210 milímetros.

a) Determine la media y la desviación estándar de la longitud de los cables.

b) Si las especificaciones para la longitud son $1\,195 < x < 1\,205$ milímetros, ¿qué proporción de los cables cumplen con las especificaciones?

5-5 DISTRIBUCIÓN CONTINUA UNIFORME

Con frecuencia se presentan diversas distribuciones continuas en las aplicaciones. En las secciones restantes de este capítulo se describen estas distribuciones y se incluyen ejemplos de cálculos de probabilidades, medias y varianzas.

La distribución continua más simple es análoga a su contraparte discreta.

Definición

Una variable aleatoria continua X con función de densidad de probabilidad

$$f(x) = 1/(b-a), \quad a \le x \le b \tag{5-6}$$

tiene una **distribución continua uniforme**.

En la figura 5-8 se ilustra la función de densidad de probabilidad de una variable aleatoria continua uniforme.

La media de la variable aleatoria continua uniforme X es

$$E(X) = \int_a^b \frac{x}{b-a}\,dx = \left.\frac{0.5x^2}{b-a}\right|_a^b = (a+b)/2$$

La varianza de X es

$$V(X) = \int_a^b \frac{\left(x - \left(\frac{a+b}{2}\right)\right)^2}{b-a}\,dx = \left.\frac{\left(x - \frac{a+b}{2}\right)^3}{3(b-a)}\right|_a^b = (b-a)^2/12$$

Estos resultados se resumen como sigue:

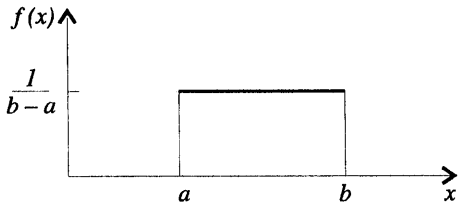


Figura 5-8 Función de densidad de probabilidad continua uniforme.

La media y la varianza de una variable aleatoria continua uniforme X en $a \leq x \leq b$ son

$$\mu = E(X) = (a+b)/2 \quad \text{y} \quad \sigma^2 = V(X) = (b-a)^2/12 \tag{5-7}$$

EJEMPLO 5-8

Sea que la variable aleatoria continua X denote la corriente medida en un alambre delgado de cobre en miliamperes. Suponga que el rango de X es [0, 20 mA], y suponga que la función de densidad de probabilidad de X es $f(x) = 0.05$, $0 \leq x \leq 20$.

¿Cuál es la probabilidad de que una medición de la corriente esté entre 5 y 10 miliamperes? El área sombreada de la figura 5-9 indica la probabilidad pedida.

$$P(5 < X < 10)) = \int_5^{10} f(x)\,dx$$
$$= 5(0.05)) = 0.25$$

Además,

$$E(X) = 10 \text{ mA} \quad \text{y} \quad V(X) = 20^2/12 = 33.33 \text{ mA}^2$$

Por consiguiente, la desviación estándar de X es 5.77 mA.

La función de distribución acumulada de una variable aleatoria continua uniforme se obtiene por integración. Si $a < x < b$, entonces

$$F(x) = \int_a^x 1/(b-a)\,du$$
$$= x/(b-a) - a/(b-a)$$

Por lo tanto, la descripción completa de la función de distribución acumulada de una variable aleatoria continua uniforme es

$$F(x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \leq x < b \\ 1 & b \leq x \end{cases}$$

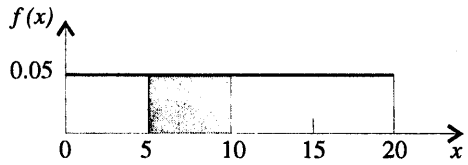


Figura 5-9 Probabilidad del ejemplo 5-8.

EJERCICIOS DE LA SECCIÓN 5-5

5-28. Suponga que X tiene una distribución continua uniforme en el intervalo [1.5, 5.5].
a) Determine la media, la varianza y la desviación estándar de X.
b) ¿Cuál es $P(X < 2.5)$?

5-29. Suponga que X tiene una distribución continua uniforme en el intervalo [–1, 1].
a) Determine la media, la varianza y la desviación estándar de X.
b) Determine el valor de x tal que $P(-x < X < x) = 0.90$.

5-30. La distribución para el peso neto en libras de un herbicida químico empacado es uniforme para $49.75 < x < 50.25$ libras.
a) Determine la media y la varianza del peso de los paquetes.
b) Determine la función de distribución acumulada del peso de los paquetes.
c) Determine $P(X < 50.1)$.

5-31. El grosor del reborde de un componente aeronáutico tiene una distribución uniforme entre 0.95 y 1.05 milímetros.
a) Determine la función de distribución acumulada del grosor de los rebordes.
b) Determine la proporción de los rebordes que exceden 1.02 milímetros.
c) ¿Qué grosor exceden 90% de los rebordes?
d) Determine la media y la varianza del grosor de los rebordes.

5-32. El espesor del recubrimiento fotoprotector aplicado a las obleas en la fabricación de semiconductores en un sitio particular de la oblea tiene una distribución uniforme entre 0.2050 y 0.2150 micrones (milésimas de milímetro).
a) Determine la función de distribución acumulada del espesor del recubrimiento fotoprotector.
b) Determine la proporción de obleas cuyo espesor del recubrimiento fotoprotector excede 0.2125 micrones (milésimas de milímetro).
c) ¿Qué espesor exceden 10% de las obleas?
d) Determine la media y la varianza del espesor del recubrimiento fotoprotector.

5-33. La función de densidad de probabilidad del tiempo necesario para terminar una operación de ensamblaje es $f(x) = 0.1$ para $30 < x < 40$ segundos.
a) Determine la proporción de unidades ensambladas cuya terminación necesita más de 35 segundos.
b) ¿Qué tiempo es excedido por 90% de las unidades ensambladas?
c) Determine la media y la varianza del tiempo de ensamblaje.

5-6 DISTRIBUCIÓN NORMAL

Sin lugar a dudas, el modelo de uso más generalizado para la distribución de una variable aleatoria es la **distribución normal**. En el capítulo 2 se muestran varios histogramas con formas de campana similares. Un resultado fundamental, conocido como el **teorema del límite central**, implica que los histogramas con frecuencia tienen esta forma característica, al menos aproximadamente. Siempre que se hace la repetición de un experimento aleatorio, la variable aleatoria que es igual al resultado promedio (o total) en las repeticiones tiende a tener una distribución normal, cuando el número de repeticiones es grande. De Moivre presentó este resultado en 1733. Desafortunadamente, su trabajo estuvo perdido durante algún tiempo, y Gauss desarrolló de manera independiente la distribución normal casi 100 años después. Aun cuando posteriormente se le dio el crédito a De Moivre por la deducción, una distribución normal también se conoce como distribución **de Gauss o gaussiana**.

¿Cuándo se saca el promedio (o el total) de los resultados? Casi siempre. En el ejemplo 1-1, el promedio de las ocho mediciones de la fuerza de desconexión se calcularon como 13.0 libras-

pie. Si se supone que cada medición es el resultado de una repetición de un experimento aleatorio, entonces puede usarse la distribución normal para sacar conclusiones aproximadas de este promedio. Estas conclusiones son el tema central de los capítulos subsecuentes del presente libro.

Por otra parte, en ocasiones el teorema del límite central es menos obvio. Por ejemplo, supongamos que la desviación (o error) en la longitud de una pieza maquinada es la suma de un gran número de efectos infinitesimales, tales como variaciones de la temperatura y la humedad, vibraciones, variaciones en el ángulo de corte, desgaste de la herramienta de corte, desgaste de los cojinetes, variaciones en la rapidez de rotación, variaciones de montaje e instalación, variaciones en las numerosas características de materias primas, y variación en los niveles de contaminación. Si los errores que intervienen son independientes y tienen la misma factibilidad de ser positivos o negativos, entonces puede demostrarse que el error total tiene una distribución normal aproximada. Además, la distribución normal surge en el estudio de numerosos fenómenos físicos básicos. Por ejemplo, el físico Maxwell desarrolló una distribución normal a partir de supuestos sencillos para las velocidades de las moléculas.

Se hace mención de la base teórica de una distribución normal para justificar, en parte, la forma compleja de la función de densidad de probabilidad. Nuestro objetivo ahora es calcular probabilidades para una variable aleatoria normal. El establecimiento más cuidadoso del teorema del límite central se hará más adelante.

Variables aleatorias con medias y varianzas diferentes pueden modelarse por medio de funciones de densidad de probabilidad normal, con la elección adecuada del centro y la anchura de la curva. El valor de $E(X) = \mu$ determina el centro de la función de densidad de probabilidad y el valor de $V(X) = \sigma^2$ determina la anchura. En la figura 5-10 se ilustran varias funciones de densidad de probabilidad normal con valores escogidos de μ y σ^2. Todas tienen la característica curva en forma de campana, pero el centro y la dispersión varían. En la siguiente definición se da la fórmula para las funciones de densidad de probabilidad normal.

Definición

Una variable aleatoria X con función de densidad de probabilidad

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad \text{para} \quad -\infty < x < \infty \tag{5-8}$$

tiene una **distribución normal** con parámetros μ, donde $-\infty < \mu < \infty$, y $\sigma > 0$. Además,

$$E(X) = \mu \quad \text{y} \quad V(X) = \sigma^2 \tag{5-9}$$

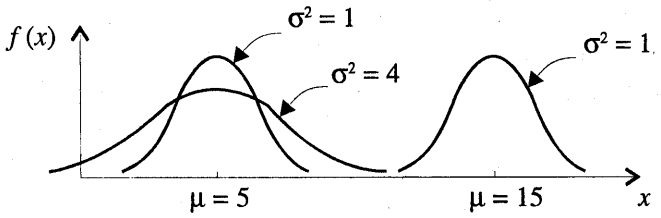


Figura 5-10 Funciones de densidad de probabilidad normal para valores escogidos de los parámetros μ y σ^2.

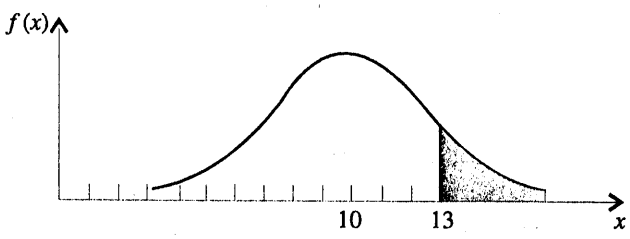


Figura 5-11 Probabilidad de $X > 13$ para una variable aleatoria normal con $\mu = 10$ y $\sigma^2 = 4$.

La notación $N(\mu, \sigma^2)$ denota una distribución normal con media μ y varianza σ^2.

EJEMPLO 5-9

Suponga que las mediciones de la corriente en una tira de alambre siguen una distribución normal con una media de 10 miliamperes y una varianza de 4 (miliamperes)^2. ¿Cuál es la probabilidad de que una medición exceda 13 miliamperes?

Sea que X denote la corriente en miliamperes. La probabilidad pedida puede representarse como $P(X > 13)$. Esta probabilidad se indica con el área sombreada bajo la función de densidad de probabilidad normal en la figura 5-11. Desafortunadamente, no hay una expresión de forma elemental para la integral de una función de densidad de probabilidad normal, y las probabilidades basadas en la distribución normal generalmente se encuentran por métodos numéricos o en una tabla (este punto se tratará más adelante).

En la figura 5-12 se resumen algunos resultados de utilidad en lo que respecta a una distribución normal. Para cualquier variable aleatoria normal,

$$P(\mu - \sigma < X < \mu + \sigma) = 0.6827$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9545$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9973$$

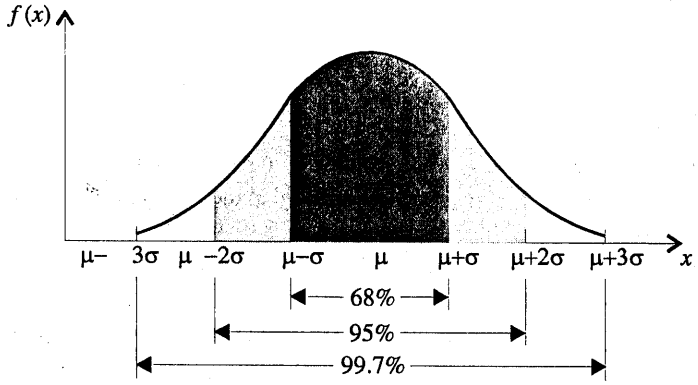


Figura 5-12 Probabilidades asociadas con una distribución normal.

Por la simetría de $f(x)$, $P(X > \mu) = P(X < \mu) = 0.5$. Puesto que $f(x)$ es positiva para toda x, este modelo asigna una probabilidad a todo intervalo de la recta real. Sin embargo, la función de densidad de probabilidad decrece conforme x se aparta más de μ. Por consiguiente, la probabilidad de que una medición esté lejos de μ es pequeña y a cierta distancia de μ la probabilidad de un intervalo puede aproximarse como cero. El área bajo la función de densidad de probabilidad normal después de 3σ de la media es muy pequeña. Este hecho resulta conveniente para gráficas aproximadas rápidas de una función de densidad de probabilidad normal. Los bosquejos de las gráficas sirven de ayuda para Determine probabilidades. Debido a que más de 0.9973 de la probabilidad de una distribución normal está dentro del intervalo $(\mu - 3\sigma, \mu + 3\sigma)$, *es común referirse a* 6σ *como la anchura de una distribución normal.* Aplicando la integración numérica es posible demostrar que el área bajo la función de densidad de probabilidad normal de $-\infty < x < \infty$ es 1.

Definición

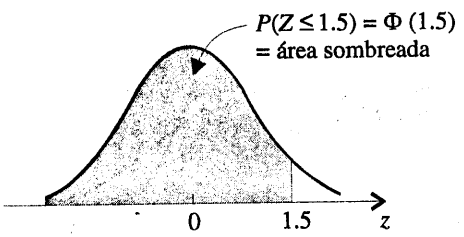
A una variable aleatoria normal con $\mu = 0$ y $\sigma^2 = 1$ se le llama **variable aleatoria normal estándar**. Una variable aleatoria normal estándar se denota como Z.

La tabla II del apéndice contiene las probabilidades acumuladas de una variable aleatoria normal estándar. El uso de la tabla II se ilustra con el ejemplo siguiente:

EJEMPLO 5-10

Suponga que Z es una variable aleatoria normal estándar. La tabla II del apéndice contiene probabilidades de la forma $P(Z \leq z)$. En la figura 5-13 se ilustra el uso de la tabla II para encontrar $P(Z \leq 1.5)$. Se lee hacia abajo la columna z hasta el renglón que es igual a 1.5. La probabilidad se lee en la columna adyacente, con el encabezado 0.00, como 0.93319.

Los encabezados de las columnas corresponden al dígito de las centésimas del valor de z en $P(Z \leq z)$. Por ejemplo, $P(Z \leq 1.53)$ se encuentra leyendo hacia abajo la columna z hasta el renglón 1.5 y después se selecciona la probabilidad en la columna con el encabezado 0.03 como 0.93699.



z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

Figura 5-13 Función de densidad de probabilidad normal estándar.

Definición

La función de distribución acumulada de una variable aleatoria normal estándar se denota como

$$\Phi(z) = P(Z \leq z)$$

Por consiguiente, la tabla II del apéndice proporciona los valores de $\Phi(z)$ para valores de Z. Las funciones de distribución acumulada para variables aleatorias normales también pueden encontrarse en gran número de paquetes de computadora. Pueden usarse de la misma manera que la tabla II del apéndice para obtener las probabilidades de estas variables aleatorias.

Las probabilidades que no son de la forma $P(Z \leq z)$ se encuentran aplicando las reglas básicas de la probabilidad y la simetría de la distribución normal junto con la tabla II del apéndice. En los ejemplos siguientes se ilustra el método.

EJEMPLO 5-11

En la figura 5-14 se muestran gráficamente los siguientes cálculos. En la práctica, una probabilidad suele redondearse a uno o dos dígitos significativos.

1) $P(Z > 1.26) = 1 - P(Z \leq 1.26) = 1 - 0.89616 = 0.10384$

2) $P(Z < -0.86) = 0.19490$

3) $P(Z > -1.37) = P(Z < 1.37) = 0.91465$

4) $P(-1.25 < Z < 0.37)$. Esta probabilidad puede encontrarse como la diferencia de dos áreas, $P(Z < 0.37) - P(Z < -1.25)$. Ahora bien,

$$P(Z < 0.37) = 0.64431 \quad \text{y} \quad P(Z < -1.25) = 0.10565$$

Por lo tanto,

$$P(-1.25 < Z < 0.37) = 0.64431 - 0.10565 = 0.53866$$

5) $P(Z \leq -4.6)$ no puede encontrarse de manera exacta en la tabla II. Sin embargo, puede usarse la última entrada de la tabla para encontrar que $P(Z \leq -3.99) = 0.00003$. Puesto que $P(Z \leq -4.6) < P(Z \leq 3.99)$, $P(Z \leq -4.6)$ es prácticamente cero.

6) Encuentre el valor de z tal que $P(Z > z) = 0.05$. Esta expresión de probabilidad puede escribirse como $P(Z \leq z) = 0.95$. Ahora se usa la tabla II en sentido inverso. Se busca en las probabilidades hasta encontrar el valor que corresponda a 0.95. La solución se ilustra en la figura 5-14. No se encuentra 0.95 exactamente; el valor más próximo es 0.95053, que corresponde a $z = 1.65$.

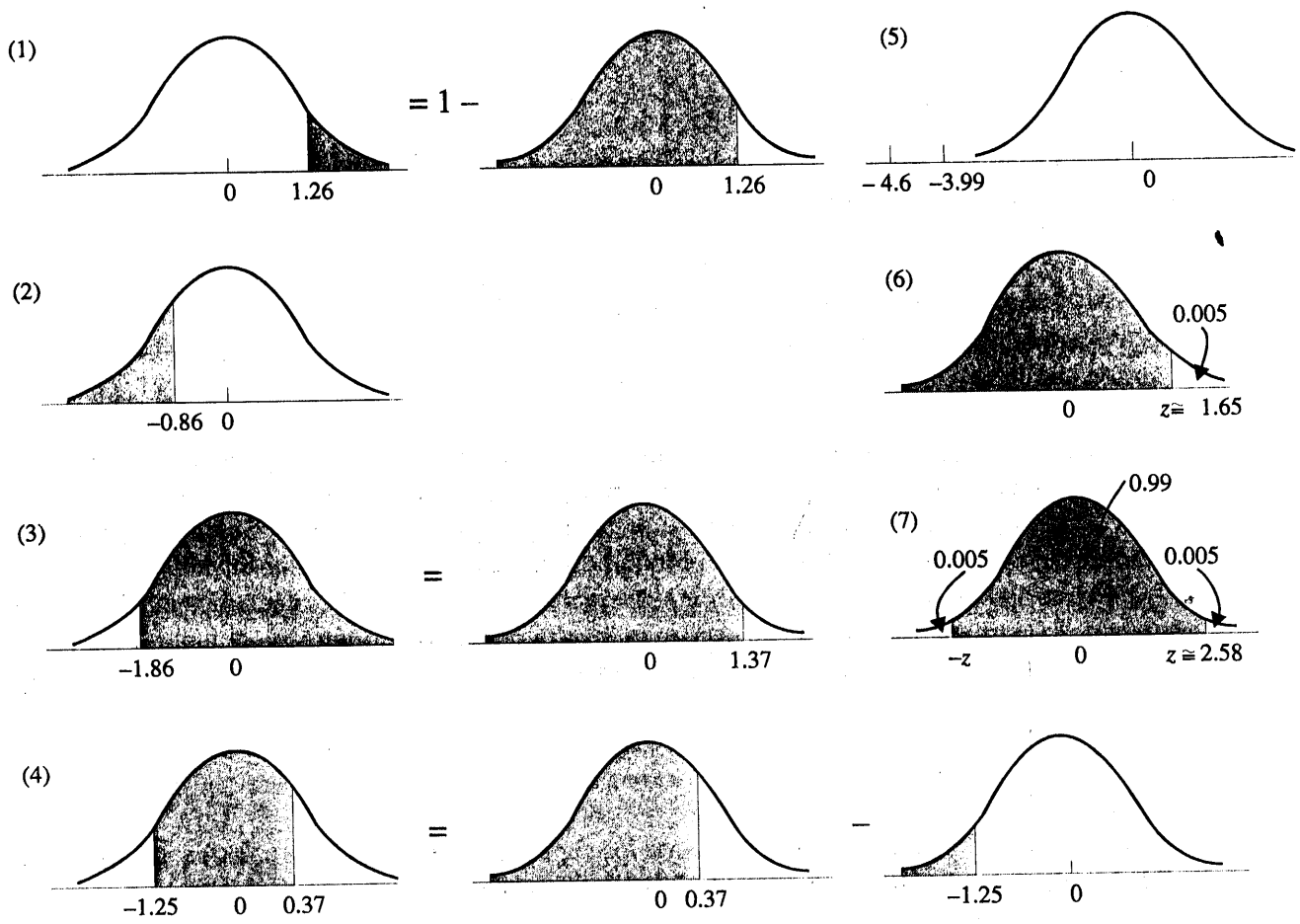


Figura 5-14 Ilustraciones gráficas de las distribuciones normales estándar.

7) Encuentre el valor de z tal que $P(-z < Z < z) = 0.99$. Debido a la simetría de la distribución normal, si el área de la región sombreada en la figura 5-14(7) es igual a 0.99, entonces el área en cada cola de la distribución debe ser igual a 0.005. Por lo tanto, el valor de z corresponde a una probabilidad de 0.995 en la tabla II. La probabilidad más próxima en la tabla II es 0.99506, cuando $z = 2.58$.

Los ejemplos precedentes muestran la forma de calcular probabilidades de variables aleatorias normales estándares. Para usar el mismo enfoque con una variable aleatoria normal arbitraria se necesitaría una tabla separada para cada posible par de valores de μ y σ. Por fortuna, todas las distribuciones de probabilidad normales se relacionan algebraicamente, y la tabla II del apéndice puede usarse para encontrar las probabilidades asociadas con una variable aleatoria normal arbitraria haciendo antes una transformación simple.

Si X es una variable aleatoria normal con $E(X) = \mu$ y $V(X) = \sigma^2$, entonces la variable aleatoria

$$Z = \frac{X - \mu}{\sigma} \tag{5-10}$$

es una variable aleatoria normal con $E(Z) = 0$ y $V(Z) = 1$. Es decir, Z es una variable aleatoria normal estándar.

A la creación de una nueva variable aleatoria por medio de esta transformación se le llama **estandarización**. La variable aleatoria Z representa la distancia de X de su media en términos de desviaciones estándar. Es el paso clave para calcular una probabilidad con una variable aleatoria normal arbitraria.

EJEMPLO 5-12

Suponga que las mediciones de la corriente en una tira de alambre siguen una distribución normal con una media de 10 miliamperes y una varianza de 4 (miliamperes)^2. ¿Cuál es la probabilidad de que una medición exceda 13 miliamperes?

Sea que X denote la corriente en miliamperes. La probabilidad pedida puede representarse como $P(X > 13)$. Sea $Z = (X - 10)/2$. La relación entre los diferentes valores de X y los valores transformados de Z se muestran en la figura 5-15. Se observa que $X > 13$ corresponde a $Z > 1.5$. Por lo tanto, por la tabla II,

$$P(X > 13) = P(Z > 1.5) = 1 - P(Z \le 1.5) = 1 - 0.93319 = 0.06681$$

En lugar de utilizar la figura 5-15, la probabilidad también pudo haberse calculado a partir de la desigualdad $X > 13$. Es decir,

$$P(X > 13) = P((X - 10)/2 > (13 - 10)/2) = P(Z > 1.5) = 0.06681$$

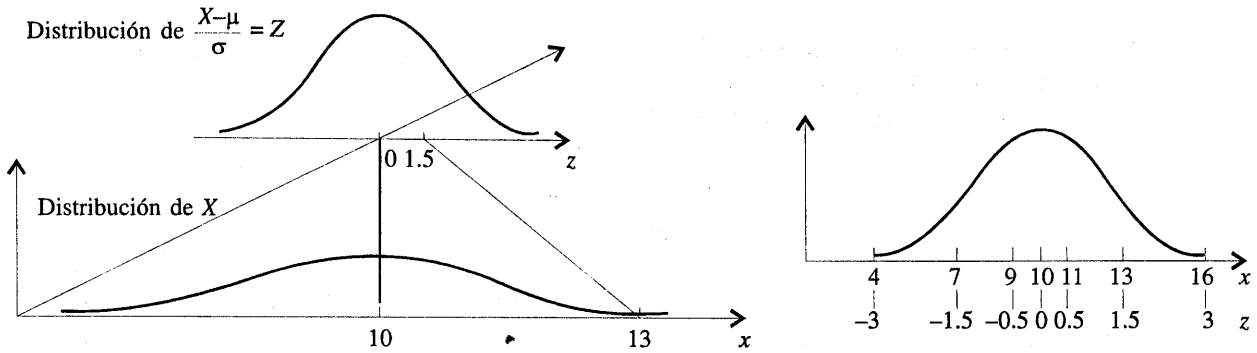


Figura 5-15 Estandarización de una variable aleatoria normal.

En el ejemplo anterior, el valor 13 se transforma en 1.5 mediante la estandarización; y se llama con frecuencia a 1.5 el **valor** z asociado con una probabilidad.

En el recuadro siguiente se resume el cálculo de probabilidades derivadas de variables aleatorias normales.

Suponga que X es una variable aleatoria normal con media μ y varianza σ^2. Entonces,

$$P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P(Z \le z) \tag{5-11}$$

donde

Z es una **variable aleatoria normal estándar**, y

$z = (x - \mu)/\sigma$ es el **valor** z que se obtiene al **estandarizar** X.

La probabilidad se obtiene introduciendo el valor $z = (x - \mu)/\sigma$ en la **tabla II del apéndice**.

EJEMPLO 5-13

Continuando con el ejemplo anterior, ¿cuál es la probabilidad de que una medición de la corriente esté entre 9 y 11 miliamperes? Por la figura 5-15, o procediendo algebraicamente, se tiene

$$\begin{aligned} P(9 < X < 11) &= P((9 - 10)/2 < (X - 10)/2 < (11 - 10)/2) \\ &= P(-0.5 < Z < 0.5) \\ &= P(Z < 0.5) - P(Z < -0.5) \\ &= 0.69146 - 0.30854 \\ &= 0.38292 \end{aligned}$$

Determine el valor para el que la probabilidad de que una medición de la corriente esté abajo de este valor sea 0.98. El valor pedido se ilustra gráficamente en la figura 5-16. Se necesita el valor de x tal que $P(X < x) = 0.98$. Después de hacer la estandarización, esta expresión de probabilidad puede escribirse como

$$\begin{aligned} P(X < x) &= P((X - 10)/2 < (x - 10)/2) \\ &= P(Z < (x - 10)/2) \\ &= 0.98 \end{aligned}$$

La tabla II se usa para encontrar el valor z tal que $P(Z < z) = 0.98$. La probabilidad más próxima de la tabla II da como resultado

$$P(X < 2.05) = 0.97982$$

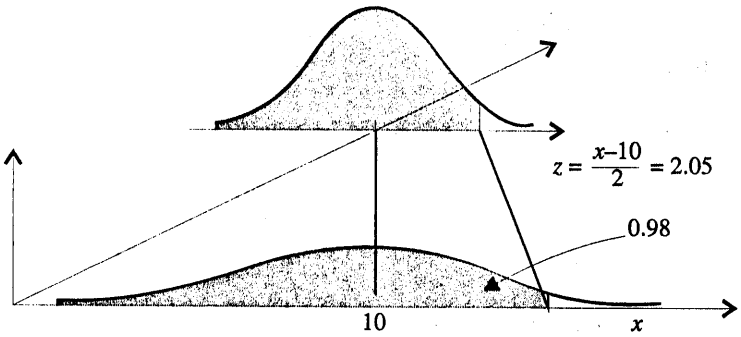


Figura 5-16 Determinación del valor de x que satisface una probabilidad dada.

Por lo tanto, $(x - 10)/2 = 2.05$, y la transformación para estandarizar se usa a la inversa para despejar x. El resultado es

$$x = 2(2.05) + 10 = 14.1 \text{ miliamperes}$$

EJEMPLO 5-14

Suponga que en la detección de una señal digital el ruido de fondo sigue una distribución normal con una media de 0 volts y una desviación estándar de 0.45 volts. Si el sistema asume que se ha transmitido un uno digital cuando el voltaje excede 0.9, ¿cuál es la probabilidad de detectar un uno digital cuando no se envió?

Sea que la variable aleatoria N denote el voltaje del ruido. La probabilidad pedida es

$$P(N > 0.9) = P\left(\frac{N}{0.45} > \frac{0.9}{0.45}\right) = P(Z > 2) = 1 - 0.97725 = 0.2275$$

Esta probabilidad puede describirse como la probabilidad de una detección falsa.

Determine los límites simétricos alrededor de 0 que incluyan al 99% de todas las lecturas de ruido. El problema requiere encontrar x tal que $P(-x < N < x) = 0.99$. En la figura 5-17 se muestra una gráfica. Ahora bien,

$$\begin{aligned} P(-x < N < x) &= P(-x/0.45 < N/0.45 < x/0.45) \\ &= P(-x/0.45 < Z < x/0.45) = 0.99 \end{aligned}$$

Por la tabla II

$$P(-2.58 < Z < 2.58) = 0.99$$

Por lo tanto,

$$x/0.45 = 2.58$$

y

$$x = 2.58(0.45) = 1.16$$

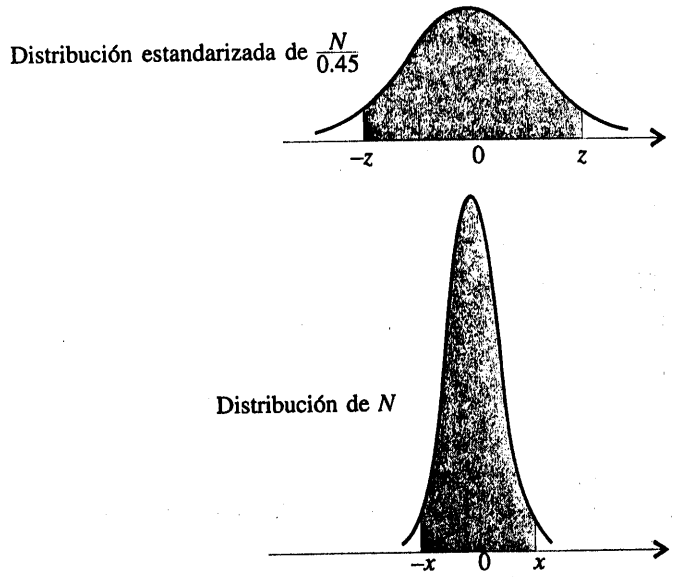


Figura 5-17 Determinación del valor de x que satisface una probabilidad dada.

Suponga que un uno digital se representa como un corrimiento en la media de la distribución del ruido a 1.8 volts. ¿Cuál es la probabilidad de que no se detecte un uno digital? Sea que la variable aleatoria S denote el voltaje cuando se transmite un uno digital. Entonces,

$$\begin{aligned} P(S<0.9) &= P\left(\frac{S-1.8}{0.45} < \frac{0.9-1.8}{0.45}\right) \\ &= P(Z<-2) \\ &= 0.02275 \end{aligned}$$

Esta probabilidad puede interpretarse como la probabilidad de una señal que se pierde.

EJEMPLO 5-15

El diámetro de un eje en un propulsor de almacenamiento óptico, tiene una distribución normal con una media de 0.2508 pulgadas y una desviación estándar de 0.0005 pulgadas. Las especificaciones de los ejes son 0.2500 ± 0.0015 pulgadas. ¿Qué proporción de los ejes cumple con las especificaciones?

Sea que X denote el diámetro del eje en pulgadas. La probabilidad pedida se muestra en la figura 5-18 y

$$\begin{aligned} P(0.2485<X<0.2515) &= P\left(\frac{0.2485-0.2508}{0.0005} < Z < \frac{0.2515-0.2508}{0.0005}\right) \\ &= P(-4.6<Z<1.4) \\ &= P(Z<1.4)-P(Z<-4.6) \\ &= 0.91924-0.0000 \\ &= 0.91924 \end{aligned}$$

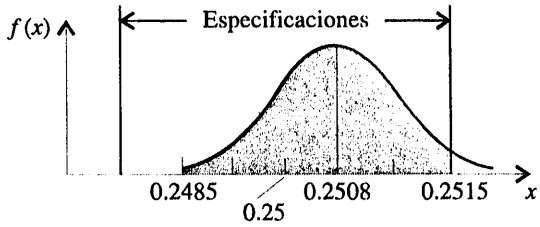


Figura 5-18 Distribución del ejemplo 5-15.

La mayoría de los ejes que no cumplen son demasiado grandes, debido a que la media del proceso se localiza muy cerca del límite superior de la especificación. Si el proceso se centra de tal modo que la media del proceso sea igual al valor especificado de 0.2500, entonces

$$\begin{aligned} P(0.2485 < X < 0.2515) &= P\left(\frac{0.2485-0.2500}{0.0005} < Z < \frac{0.2515-0.2500}{0.0005}\right) \\ &= P(-3 < Z < 13) \\ &= P(Z < 3) - P(Z < -3) \\ &= 0.99865 - 0.00135 \\ &= 0.9973 \end{aligned}$$

Al centrar nuevamente el proceso, el producto aumenta a aproximadamente 99.73%.

En el desarrollo siguiente se demuestra que la media y la varianza de una variable aleatoria normal son μ y σ^2, respectivamente. La media de x es

$$E(X) = \int_{-\infty}^{\infty} x \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma} dx$$

Al hacer el cambio de variable $y = (x - \mu)/\sigma$, la integral se vuelve

$$E(X) = \mu \int_{-\infty}^{\infty} x \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy + \sigma \int_{-\infty}^{\infty} y \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy$$

La primera integral de esta expresión es igual a 1 porque $\frac{e^{-y^2/2}}{\sqrt{2\pi}}$ es una función de densidad de probabilidad, y se encuentra que la segunda integral es 0, bien al hacer el cambio de variable $u = -y^2/2$ o al observar la simetría del integrando respecto de $y = 0$. Por lo tanto, $E(X) = \mu$.

La varianza de X es

$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma} dx$$

Al hacer el cambio de variable $y = (x - m)/\sigma$, la integral se vuelve

$$V(X)=\sigma^2 \int_{-\infty}^{\infty} y^2 \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy$$

Al integrar por partes con $u = y$ y $dv = y\frac{e^{-y^2/2}}{\sqrt{2\pi}} dy$, se encuentra que $V(X)$ es σ^2.

EJERCICIOS DE LA SECCIÓN 5-6

5-34. Use la tabla II del apéndice para determinar las siguientes probabilidades de la variable aleatoria normal estándar Z.

a) $P(Z < 1.32)$
b) $P(Z < 3.0)$
c) $P(Z > 1.45)$
d) $P(Z > -2.15)$
e) $P(-2.34 < Z < 1.76)$

5-35. Use la tabla II del apéndice para determinar las siguientes probabilidades de la variable aleatoria normal estándar Z.

a) $P(-1 < Z < 1)$
b) $P(-2 < Z < 2)$
c) $P(-3 < Z < 3)$
d) $P(Z > 3)$
e) $P(0 < Z < 1)$

5-36. Suponga que Z tiene una distribución normal estándar. Use la tabla II del apéndice para determinar el valor de z que resuelve cada una de las expresiones siguientes:

a) $P(Z < z) = 0.9$
b) $P(Z < z) = 0.5$
c) $P(Z > z) = 0.1$
d) $P(Z > z) = 0.9$
e) $P(-1.24 < Z < z) = 0.8$

5-37. Suponga que Z tiene una distribución normal estándar. Use la tabla II del apéndice para determinar los valores para z que resuelve cada una de las expresiones siguientes:

a) $P(-z < Z < z) = 0.95$
b) $P(-z < Z < z) = 0.99$
c) $P(-z < Z < z) = 0.68$
d) $P(-z < Z < z) = 0.9973$

5-38. Suponga que X tiene una distribución normal con una media de 10 y una desviación estándar de 2. Determine lo siguiente:

a) $P(X < 13)$
b) $P(X > 9)$
c) $P(6 < X < 14)$
d) $P(2 < X < 4)$
e) $P(-2 < X < 8)$

5-39. Suponga que X tiene una distribución normal con una media de 10 y una desviación estándar de 2. Determine el valor de x que resuelve cada una de las expresiones siguientes:

a) $P(X > x) = 0.5$
b) $P(X > x) = 0.95$
c) $P(x < X < 10) = 0.2$
d) $P(-x < X - 10 < x) = 0.95$
e) $P(-x < X - 10 < x) = 0.99$

5-40. Suponga que X tiene una distribución normal con una media de 5 y una desviación estándar de 4. Determine lo siguiente:

a) $P(X < 11)$
b) $P(X > 0)$
c) $P(3 < X < 7)$
d) $P(-2 < X < 9)$
e) $P(2 < X < 8)$

5-41. Suponga que X tiene una distribución normal con una media de 5 y una desviación estándar de 4. Determine el valor de x que resuelve cada una de las expresiones siguientes:

a) $P(X > x) = 0.5$
b) $P(X > x) = 0.95$
c) $P(x < X < 9) = 0.2$
d) $P(3 < X < x) = 0.95$
e) $P(-x < X < x) = 0.99$

5-42. La resistencia a la compresión de muestras de cemento puede modelarse mediante una distribución normal con una media de 6 000 kilogramos por centímetro cuadrado y una

desviación estándar de 100 kilogramos por centímetro cuadrado.

a) ¿Cuál es la probabilidad de que la resistencia a la compresión de una muestra sea menor que 6 250 kg/cm^2?

b) ¿Cuál es la probabilidad de que la resistencia a la compresión de una muestra esté entre 5 800 y 5 900 kg/cm^2?

c) ¿Cuál es la resistencia a la compresión que excede 95% de las muestras?

5-43. La resistencia a la tensión del papel se modela mediante una distribución normal con una media de 35 libras por pulgada cuadrada y una desviación estándar de 2 libras por pulgada cuadrada.

a) ¿Cuál es la probabilidad de que la resistencia a la tensión de una muestra sea menor que 40 lb/pulg^2?

b) Si las especificaciones requieren que la resistencia a la tensión exceda 30 lb/pulg^2, ¿qué proporción de las muestras se desecha?

5-44. Se supone que la anchura de las rayas espectrales de una herramienta usada para fabricar semiconductores tiene una distribución normal con una media de 0.5 micrones (milésimas de milímetro) y una desviación estándar de 0.05 micrones (milésimas de milímetro).

a) ¿Cuál es la probabilidad de que la anchura de una raya espectral sea mayor que 0.62 micrones (milésimas de milímetro)?

b) ¿Cuál es la probabilidad de que la anchura de una raya espectral esté entre 0.47 y 0.63 micrones (milésimas de milímetro)?

c) ¿Abajo de qué valor está el 90% de las anchuras de las rayas espectrales?

5-45. El volumen de llenado de una máquina automatizada usada para llenar latas de una bebida carbonatada tiene una distribución normal con una media de 12.4 onzas líquidas y una desviación estándar de 0.1 onzas líquidas.

a) ¿Cuál es la probabilidad de que un volumen de llenado sea menor que 12 onzas líquidas?

b) Si se desechan todas las latas con menos de 12.1 onzas o con más de 12.6 onzas, ¿qué proporción de las latas se desecha?

c) Determine las especificaciones simétricas alrededor de la media que incluyan al 99% de las latas.

5-46. En el ejercicio anterior, suponga que la media de la operación de llenado puede ajustarse con facilidad, pero la desviación estándar se mantiene en 0.1 onzas.

a) ¿En qué valor deberá fijarse la media para que 99.9% de las latas excedan 12 onzas?

b) ¿En qué valor deberá fijarse la media para que 99.9% de las latas excedan 12 onzas si la desviación estándar puede reducirse a 0.05 onzas líquidas?

5-47. El tiempo de reacción de un conductor a un estímulo visual tiene una distribución normal con una media de 0.4 segundos y una desviación estándar de 0.05 segundos.

a) ¿Cuál es la probabilidad de que una reacción requiera más de 0.5 segundos?

b) ¿Cuál es la probabilidad de que una reacción requiera entre 0.4 y 0.5 segundos?

c) ¿Cuál es el tiempo de reacción que se excede 90% de las veces?

5-48. La longitud de un estuche de plástico moldeado por inyección para mantener cinta magnética tiene una distribución normal con una longitud media de 90.2 milímetros y una desviación estándar de 0.1 milímetros.

a) ¿Cuál es la probabilidad de que la longitud de un estuche sea mayor que 90.3 milímetros o menor que 89.7 milímetros?

b) ¿En qué valor deberá fijarse la media del proceso para obtener el mayor número de estuches con longitud entre 89.7 y 90.3 milímetros?

c) Si los estuches que no midan entre 89.7 y 90.3 milímetros se desechan, ¿cuál es el resultado para la media del proceso que se seleccionó en el inciso b)?

5-49. En el ejercicio anterior, suponga que el centro del proceso se establece de tal forma que la media es 90 milímetros y la desviación estándar es 0.1 milímetros. Suponga que se miden 10 estuches y se supone que son independientes.

a) ¿Cuál es la probabilidad de que los 10 estuches midan entre 89.7 y 90.3 milímetros?

b) ¿Cuál es el número esperado de los 10 estuches que miden entre 89.7 y 90.3 milímetros?

5-50. El tiempo de incapacidad por enfermedad de los empleados de una empresa en un mes, tiene una distribución normal con una media de 100 horas y una desviación estándar de 20 horas.

a) ¿Cuál es la probabilidad de que el tiempo de incapacidad del mes próximo esté entre 50 y 80 horas?

b) ¿Cuánto tiempo deberá proyectarse para incapacidades si esta cantidad deberá ser excedida con una probabilidad de sólo 10%?

5-51. La vida de un láser de semiconductores con una alimentación de energía constante tiene una distribución normal con una media de 7 000 horas y una desviación estándar de 600 horas.

a) ¿Cuál es la probabilidad de que un láser falle antes de 5 000 horas?

b) ¿Cuál es la vida en horas que exceden el 95% de los láser?

c) Si se usan tres láser en un producto y se supone que fallan de manera independiente, ¿cuál es la probabilidad de que los tres sigan funcionando después de 7 000 horas?

5-52. El diámetro de los puntos producidos por una impresora matricial tiene una distribución normal con un diámetro promedio de 0.002 pulgadas y una desviación estándar de 0.0004 pulgadas.

a) ¿Cuál es la probabilidad de que el diámetro de un punto exceda 0.0026 pulgadas?

b) ¿Cuál es la probabilidad de que el diámetro de un punto mida entre 0.0014 y 0.0026 pulgadas?

c) ¿Qué desviación estándar de los diámetros se necesita para que la probabilidad del inciso b) sea 0.995?

5-53. El peso de un zapato especializado para correr tiene una distribución normal con una media de 12 onzas y una desviación estándar de 0.5 onzas.

a) ¿Cuál es la probabilidad de que un zapato pese más de 13 onzas?

b) ¿Cuál debe ser la desviación estándar del peso para que el fabricante anuncie que 99.9% de sus zapatos pesan menos de 13 onzas?

5-7 GRÁFICAS DE PROBABILIDAD

¿Cómo sabemos si una distribución normal es un modelo razonable de los datos? La **gráfica de probabilidades** es un método gráfico para determinar si los datos muestrales se ajustan a una distribución hipotetizada basada en un examen visual subjetivo de los datos. El procedimiento general es muy simple y puede hacerse con rapidez. Se acostumbra hacer la graficación de probabilidades en un papel gráfico especial, conocido como **papel probabilidad**, que se ha diseñado para la distribución hipotetizada. Es posible conseguir papel probabilidad para la distribución normal, normal logarítmica, de Weibull así como para diferentes distribuciones ji-cuadrada y gamma.

Para hacer una gráfica de probabilidad, primero se ordenan las observaciones de la muestra de la menor a la mayor. Es decir, la muestra $x_1, x_2, ..., x_n$ se ordenan como $x_{(1)}, x_{(2)}, ..., x_{(n)}$, donde $x_{(1)}$ es la observación menor, $x_{(2)}$ es la segunda observación menor y así sucesivamente, con $x_{(n)}$ denotando la mayor. Después, las observaciones ordenadas $x_{(j)}$ se grafican contra sus frecuencias acumuladas observadas $(j - 0.5)/n$ en el papel probabilidad apropiado. Si la distribución

hipotetizada describe adecuadamente los datos, los puntos graficados estarán aproximadamente en una recta; si los puntos graficados se apartan significativamente de una recta, entonces el modelo hipotetizado no es el apropiado. Por lo general, determinar si la gráfica de los datos es o no una línea recta es subjetivo. El procedimiento se ilustra en el ejemplo siguiente.

EJEMPLO 5-16

Diez observaciones de la vida de servicio efectiva en minutos de baterías usadas en una computadora personal portátil son las siguientes: 176, 191, 214, 220, 205, 192, 201, 190, 183, 185. Se hipotetiza que la vida de las baterías se modela adecuadamente con una distribución normal. Para usar la graficación de probabilidades para investigar esta hipótesis, primero se ordenan las observaciones en orden ascendente y se calculan sus frecuencias acumuladas $(j - 0.5)/10$ como sigue:

j	$x(j)$	$(j-0.5)/10$
1	176	0.05
2	183	0.15
3	185	0.25
4	190	0.35
5	191	0.45
6	192	0.55
7	201	0.65
8	205	0.75
9	214	0.85
10	220	0.95

Después se grafican en papel probabilidad normal los pares de valores $x_{(j)}$ y $(j - 0.5)/10$. La gráfica se muestra en la figura 5-19. La mayoría de los papeles probabilidad normal grafican $100(j - 0.5)/n$ en la escala vertical izquierda y $100[1 - (j - 0.5)/n]$ en la escala vertical derecha, mientras que en la escala horizontal se grafica el valor de la variable. Se ha trazado una línea recta, elegida subjetivamente, a través de los puntos graficados. Al trazar la recta, uno deberá

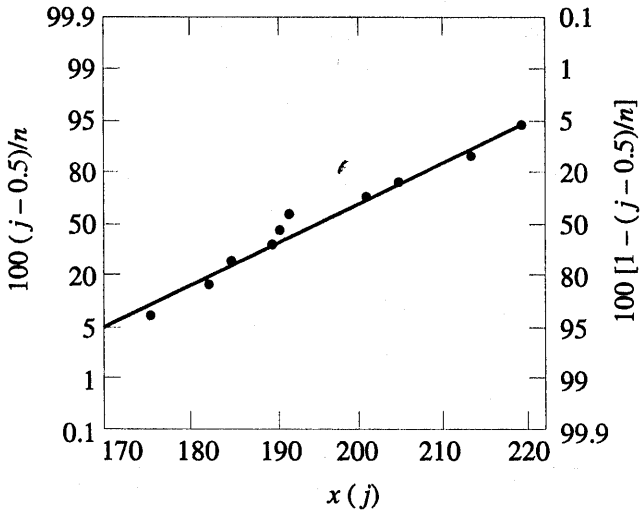


Figura 5-19 Gráfica de probabilidad normal para la vida de las baterías.

prestar mayor atención a los puntos de en medio de la gráfica que a los puntos extremos. Una buena regla empírica es trazar la recta aproximadamente entre los puntos de los percentiles 25 y 75. Así fue como se determinó la recta de la figura 5-19. Para determinar la "cercanía" de los puntos con la recta, imagine un "lápiz grueso" sobre la recta. Si todos los puntos están cubiertos por este lápiz imaginario, entonces una distribución normal describe adecuadamente los datos. Puesto que los puntos de la figura 5-19 pasarían la prueba del "lápiz grueso", se concluye que la distribución normal es un modelo apropiado.

Una gráfica de probabilidad normal también puede construirse en papel gráfico común graficando los valores normales estandarizados z_j contra $x_{(j)}$, donde los valores normales estandarizados satisfacen

$$\frac{j-0.5}{n} = P(Z \le z_j) = \Phi(z_j)$$

Por ejemplo, si $(j - 0.5)/n = 0.05$, entonces $\Phi(z_j) = 0.05$ implica que $z_j = -1.64$. Para ilustrar el procedimiento, considérense los datos del ejemplo 5-16. En la última columna de la tabla siguiente se muestran los valores normales estandarizados.

j	$x_{(j)}$	$(j-0.5)/10$	z_j
1	176	0.05	−1.64
2	183	0.15	−1.04
3	185	0.25	−0.67
4	190	0.35	−0.39
5	191	0.45	−0.13
6	192	0.55	0.13
7	201	0.65	0.39
8	205	0.75	0.67
9	214	0.85	1.04
10	220	0.95	1.64

En la figura 5-20 se presenta la gráfica de z_j contra $x_{(j)}$. La gráfica de probabilidad normal es equivalente a la de la figura 5-19.

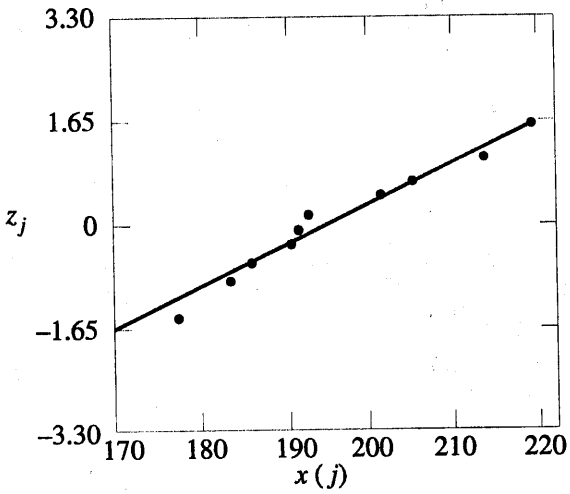


Figura 5-20 Gráfica de probabilidad normal obtenida a partir de los valores normales esandarizados.

Una aplicación muy importante de la graficación de probabilidades normales es en la *verificación de supuestos* cuando se emplean procedimientos de inferencia estadística que requieren el supuesto de normalidad.

Aquí sólo se muestran gráficas de probabilidad para una distribución normal, pero el método es igual de sencillo para otras distribuciones. Los paquetes de programas de computadora permiten generar gráficas de probabilidad para determinar muchas otras distribuciones. Una desventaja es que el método es subjetivo. En el capítulo 8 se presenta un enfoque más formal para determinar una distribución con base en la prueba de la bondad del ajuste.

EJERCICIOS DE LA SECCIÓN 5-7

5-54. Una embotelladora de gaseosas está estudiando la resistencia a la presión interna de botellas de 1 litro. Se prueba una muestra aleatoria de 16 botellas y se registran las resistencias a la presión. Los datos se muestran abajo. Grafique estos datos sobre papel probabilidad normal. ¿Sería razonable concluir que la resistencia a la presión tiene una distribución normal?

226.16 psi	211.14 psi
202.20	203.62
219.54	188.12
193.73	224.39
208.15	221.31
195.45	204.55
193.71	202.21
200.81	201.63

5-55. Se seleccionan muestras de 20 piezas de dos máquinas, y se mide una dimensión crítica en cada pieza. Los datos se muestran abajo. Grafique los datos sobre papel probabilidad normal. ¿Esta dimensión parece tener una distribución normal? ¿Qué conclusiones tentativas podrían hacerse acerca de las dos máquinas?

Máquina 1			
99.4	101.5	102.3	96.7
99.1	103.8	100.4	100.9
99.0	99.6	102.5	96.5
98.9	99.4	99.7	103.1
99.6	104.6	101.6	96.8

Máquina 2			
90.9	100.7	95.0	98.8
99.6	105.5	92.3	115.5
105.9	104.0	109.5	87.1
91.2	96.5	96.2	109.8
92.8	106.7	97.6	106.5

5-56. Después de examinar los datos de las dos máquinas del ejercicio anterior, el ingeniero del proceso concluye que la máquina 2 tiene una variabilidad mayor entre una pieza y otra. Hace algunos ajustes en la máquina que deberían reducir la variabilidad, y obtiene otra muestra de 20 partes. Las mediciones de estas piezas se muestran abajo. Grafique estos datos sobre papel probabilidad normal y haga la comparación con la gráfica de probabilidad normal del ejercicio anterior. ¿Cuáles serían sus conclusiones?

103.4	107.0	107.7	104.5
108.1	101.5	106.2	106.6
103.1	104.1	106.3	105.6
108.2	106.9	107.8	103.7
103.9	103.3	107.4	102.6

5-57. Se prueba la vida de un componente electrónico bajo condiciones de alta temperatura para acelerar el mecanismo de falla. El tiempo de falla (en horas) para 20 componentes seleccionados al azar se presenta abajo. Grafique los datos sobre papel probabilidad normal. ¿Parece factible que el tiempo de falla tenga una distribución normal?

176.1	35.3	124.5	90.6	99.6
150.4	55.0	34.9	46.0	40.4
79.6	24.7	155.7	2.42	131.5
197.6	73.0	122.8	133.8	40.4

5-8 APROXIMACIÓN NORMAL A LA DISTRIBUCIÓN BINOMIAL Y DE POISSON

La sección sobre la distribución normal se inició con un desarrollo de la distribución normal como una aproximación de una variable aleatoria binomial para un gran número de ensayos. En consecuencia, no debería ser motivo de sorpresa utilizar la distribución normal para aproximar probabilidades binomiales para casos en los que n es grande. En el ejemplo siguiente se ilustra que para muchos sistemas físicos el modelo binomial es apropiado para un valor en extremo grande de n. En estos casos, es complicado calcular probabilidades utilizando la distribución binomial. Por fortuna, la aproximación normal es más conveniente en estos casos. En la figura 5-21 se presenta un ejemplo. El área de cada rectángulo es igual a la probabilidad binomial de x. El área de los rectángulos puede aproximarse con áreas bajo la función de densidad normal.

EJEMPLO 5-17

En un canal de comunicación digital, suponga que el número de bits recibidos con error puede modelarse con una variable aleatoria binomial, y suponga que la probabilidad de que un bit se reciba con error es 1×10^{-5}. Si se transmiten 16 millones de bits, ¿cuál es la probabilidad de que más de 150 presenten errores?

Sea que la variable aleatoria X denote el número de errores. Entonces X es una variable aleatoria binomial y

$$P(X>150)=1-P(X\leq 150)$$
$$=1-\sum_{x=0}^{150}\binom{16\,000\,000}{x}(10^{-5})^{x}(1-10^{-5})^{16\,000\,000-x}$$

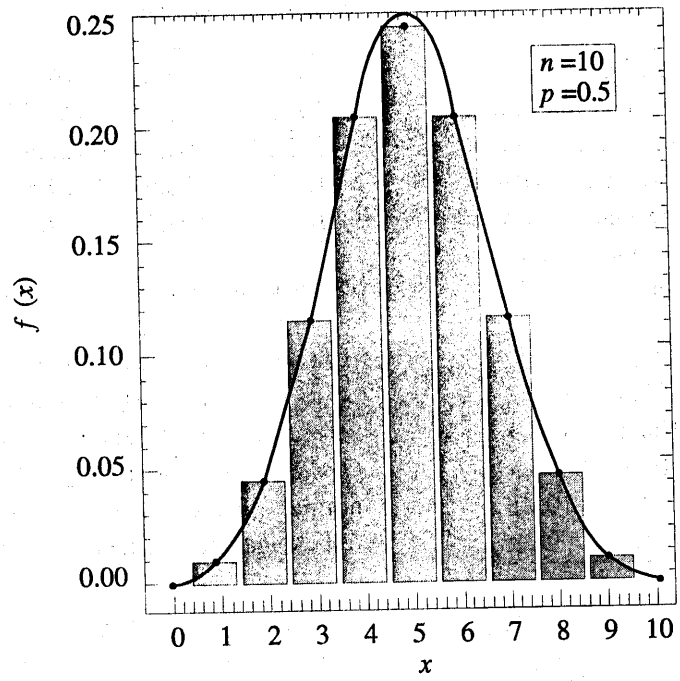


Figura 5-21 Aproximación normal de la distribución binomial.

Evidentemente, el cálculo de la probabilidad del ejemplo 5-17 es difícil. Por fortuna, puede usarse la distribución normal para obtener una excelente aproximación en este ejemplo.

Si X es una variable aleatoria binomial, entonces

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \tag{5-12}$$

es aproximadamente una variable aleatoria normal estándar. La aproximación es buena para

$$np > 5 \quad \text{y} \quad n(1-p) > 5$$

Recuérdese que para una variable binomial X, $E(X) = np$ y $V(X) = np(1-p)$. Por consiguiente, la expresión de la ecuación 5-12 no es sino la fórmula para estandarizar la variable aleatoria X. Las probabilidades en las que está presente X pueden aproximarse utilizando una distribución normal estándar. La aproximación es buena cuando n es grande en comparación con p.

EJEMPLO 5-18

El problema del canal de comunicación digital del ejemplo anterior se resuelve como sigue:

$$\begin{aligned} P(X > 150) &= P\left(\frac{x-160}{\sqrt{160(1-10^{-5})}} > \frac{150-160}{\sqrt{160(1-10^{-5})}}\right) \\ &= P(Z > -0.79) = P(Z < 0.79) = 0.785 \end{aligned}$$

Debido a que $np = (16 \times 10^6)(1 \times 10^{-5}) = 160$ y $n(1-p)$ es mucho mayor, es de esperarse que la aproximación funcione bien en este caso.

EJEMPLO 5-19

Considérese de nuevo la transmisión de bits del ejemplo 5-17. Para juzgar qué tan bien funciona la aproximación normal, suponga que sólo van a transmitirse $n = 50$ bits y que la probabilidad de un error es $p = 0.1$. La probabilidad exacta de que ocurran 2 o menos errores es

$$P(X \le 2) = \binom{50}{0} 0.9^{50} + \binom{50}{1} 0.1(0.9^{49}) + \binom{50}{2} 0.1^2 (0.9^{48})$$

Con base en la aproximación normal,

$$P(X \le 2) = P\left(\frac{X-5}{2.12} < \frac{2-5}{2.12}\right) = P(Z < -1.415) = 0.08$$

Incluso para una muestra de apenas 50 bits, la aproximación normal es razonable.

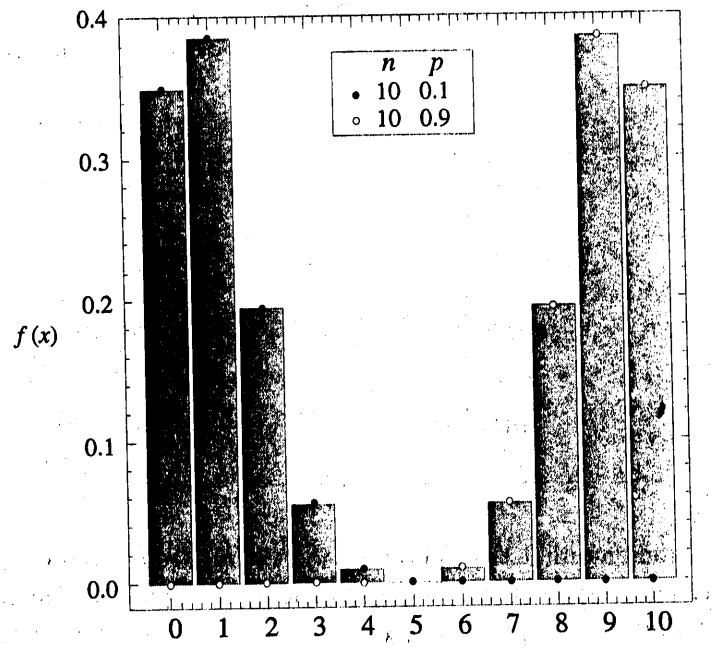


Figura 5-22 *a*) La distribución binomial no es simétrica cuando p está cerca de 0 o 1. *b*) Distribución binomial con $n = 10$, $p = 0.1$. *c*) Distribución binomial con $n = 10$, $p = 0.9$.

Puede utilizarse un factor de corrección para hacer aún mejor la aproximación. (Esto se expone en los ejercicios del final de esta sección.) Sin embargo, si np o $n(1 - p)$ son pequeños, la distribución binomial presenta un sesgo importante y la distribución normal simétrica no es una buena aproximación. En la figura 5-22 se ilustran dos casos.

Recuérdese que la distribución binomial es una aproximación satisfactoria de la distribución hipergeométrica cuando n, el tamaño de la muestra, es pequeño en comparación con N, el tamaño de la población de la que se selecciona la muestra. Una regla empírica es que la aproximación binomial es efectiva si $n/N < 0.1$. Recuérdese que para una distribución hipergeométrica, p se define como $p = K/N$. Es decir, p se interpreta como el número de éxitos en la población. Por lo tanto, la distribución normal puede proporcionar una aproximación eficaz de las probabilidades hipergeométricas cuando $n/N < 0.1$, $np > 5$ y $n(1 - p) > 5$. En la figura 5-23 se presenta un resumen de estos criterios.

Recuérdese que la distribución de Poisson se desarrolló como el límite de una distribución binomial cuando el número de ensayos se incrementa a infinito. Por consiguiente, no debe sorprender encontrar que la distribución normal puede usarse para aproximar probabilidades de una variable aleatoria de Poisson.

Distribución hipergeométrica	$\approx$ $\frac{n}{N} < 0.1$	Distribución binomial	$\approx$ $np > 5$ $n(1-p) > 5$	Distribución normal

Figura 5-23 Condiciones para aproximar probabilidades hipergeométricas y binomiales.

Si X es una variable aleatoria de Poisson con $E(X) = \lambda$ y $V(X) = \lambda$, entonces

$$Z = \frac{X - \lambda}{\sqrt{\lambda}} \tag{5-13}$$

es aproximadamente una variable aleatoria normal estándar. La aproximación es buena para

$$\lambda > 5$$

EJEMPLO 5-20

Suponga que el número de partículas de asbesto en un centímetro cuadrado de polvo sigue una distribución de Poisson con una media de 1 000. Si se analiza un centímetro cuadrado de polvo, ¿cuál es la probabilidad de encontrar menos de 950 partículas?

La probabilidad puede expresarse exactamente como

$$P(X \leq 950) = \sum_{x=0}^{950} \frac{e^{-1\,000} x^{1\,000}}{x!}$$

La dificultad del cálculo es evidente. La probabilidad puede aproximarse como

$$P(X \leq x) = P\left(Z \leq \frac{950 - 1\,000}{\sqrt{1\,000}}\right) = P(Z \leq -1.58) = 0.057$$

EJERCICIOS DE LA SECCIÓN 5-8

5-58. Suponga que X es una variable aleatoria binomial con $n = 200$ y $p = 0.4$.

a) Aproxime la probabilidad de que X sea menor o igual que 70.

b) Aproxime la probabilidad de que X sea mayor que 70 y menor que 90.

5-59. Suponga que X es una variable aleatoria binomial con $n = 100$ y $p = 0.1$.

a) Calcule la probabilidad exacta de que X sea menor que cuatro.

b) Aproxime la probabilidad de que X sea menor que cuatro y compare con el inciso a).

c) Aproxime la probabilidad de que $8 < X < 12$.

5-60. En la fabricación de chips semiconductores se producen 2% de chips defectuosos. Suponga que los chips son independientes y que un lote contiene 1 000 chips.

a) Aproxime la probabilidad de que más de 25 chips estén defectuosos.

b) Aproxime la probabilidad de que entre 20 y 30 chips estén defectuosos.

5-61. Un proveedor embarca un lote de 1 000 conectotores eléctricos. Se selecciona una muestra de 25 conectores al azar y sin reemplazo. Suponga que el lote contiene 100 conectores defectuosos.

a) Usando una aproximación binomial, ¿cuál es la probabilidad de que no haya conectores defectuosos en la muestra?

b) Use la aproximación normal para resolver el inciso a). ¿Es satisfactoria la aproximación?

c) Resuelva de nuevo los incisos a) y b) suponiendo que el tamaño del lote es 500. ¿La aproximación normal de la probabilidad de que no haya conectores defectuosos en la muestra es satisfactoria en este caso?

5-62. Un producto electrónico de oficina contiene 200 componentes electrónicos. Suponga que la probabilidad de que cada uno de los componentes opere sin falla durante la vida útil del producto es 0.999, y suponga que los componentes fallan independientemente. Aproxime la probabilidad de que cinco o más de los 200 componentes originales fallen durante la vida útil del producto.

5-63. Suponga que el número de partículas de asbesto en una muestra de un centímetro cuadrado de polvo es una variable aleatoria de Poisson con una media de 1 000. ¿Cuál es la probabilidad de que 10 centímetros cuadrados de polvo contengan más de 10 000 partículas?

5-64. Corrección de continuidad. La aproximación normal de una probabilidad binomial en ocasiones se modifica con un factor de corrección de 0.5 que mejora la aproximación. Suponga que X es binomial con $n = 50$ y $p = 0.1$. Como X es una variable aleatoria discreta, $P(X \leq 2) = P(X \leq 2.5)$. Sin embargo, la aproximación normal de $P(X \leq 2)$ puede mejorarse aplicando la aproximación a $P(X \leq 2.5)$.

a) Aproxime $P(X \leq 2)$ calculando el valor z correspondiente a $x = 2.5$.

b) Aproxime $P(X \leq 2)$ calculando el valor z correspondiente a $x = 2$.

c) Compare los resultados de los incisos a) y b) con el valor exacto de $P(X \leq 2)$ para evaluar la efectividad de la corrección de continuidad.

d) Use la corrección de continuidad para aproximar $P(X < 10)$.

5-65. Corrección de continuidad. Suponga que X es binomial con $n = 50$ y $p = 0.1$. Como X es una variable aleatoria discreta, $P(X \geq 2) = P(X \geq 1.5)$. Sin embargo, la aproximación normal de $P(X \geq 2)$ puede mejorarse aplicando la aproximación a $P(X \geq 1.5)$. La corrección de continuidad de 0.5 se suma o se resta. Una regla fácil de recordar es que la corrección de continuidad siempre se aplica para hacer más grande la probabilidad normal de aproximación.

a) Aproxime $P(X \geq 2)$ calculando el valor z correspondiente a 1.5.

b) Aproxime $P(X \geq 2)$ calculando el valor z correspondiente a 2.

c) Compare los resultados de los incisos a) y b) con el valor exacto de $P(X \geq 2)$ para evaluar la efectividad de la corrección de continuidad.

d) Use la corrección de continuidad para aproximar $P(X > 6)$.

5-66. Corrección de continuidad. Suponga que X es binomial con $n = 50$ y $p = 0.1$. Como X es una variable aleatoria discreta, $P(2 \leq X \leq 5) = P(1.5 \leq X \leq 5.5)$. Sin embargo, la aproximación normal de $P(2 \leq X \leq 5)$ puede mejorarse aplicando la aproximación a $P(1.5 \leq X \leq 5.5)$.

a) Aproxime $P(2 \leq X \leq 5)$ calculando los valores z correspondientes a 1.5 y 5.5.

b) Aproxime $P(2 \leq X \leq 5)$ calculando los valores z correspondientes a 2 y 5.

5-67. Corrección de continuidad. Suponga que X es binomial con $n = 50$ y $p = 0.1$. Entonces, $P(X = 10) = P(10 \leq X \leq 10)$. Usando los resultados para las correcciones de continuidad, puede aproximarse $P(10 \leq X \leq 10)$ aplicando la estandarización normal a $P(9.5 \leq X \leq 10.5)$.

a) Aproxime $P(X = 10)$ calculando los valores z correspondientes a 9.5 y 10.5.

b) Aproxime $P(X = 5)$.

5-68. Corrección de continuidad. En la fabricación de chips semiconductores se producen 2% de chips defectuosos. Suponga que los chips son independientes y que un lote contiene 1 000 chips.

a) Use la corrección de continuidad para aproximar la probabilidad de que entre 20 y 30 chips del lote estén defectuosos.

b) Use la corrección de continuidad para aproximar la probabilidad de que exactamente 20 chips estén defectuosos.

c) Determine el número de chips defectuosos, x, tal que la aproximación normal de la probabilidad de obtener x chips defectuosos sea la máxima.

5-9 DISTRIBUCIÓN EXPONENCIAL

En la exposición de la distribución de Poisson, una variable aleatoria se definió como el número de imperfecciones a lo largo de un alambre de cobre. La distancia entre las imperfecciones es otra variable aleatoria que suele ser de interés. Sea que la variable aleatoria X denote la longitud desde cualquier punto inicial sobre el alambre hasta que se detecta una imperfección.

Como sería de esperar, la distribución de X puede obtenerse a partir del conocimiento de la distribución del número de imperfecciones. La clave de la relación es el concepto siguiente: la distancia hasta la primera imperfección excede 3 milímetros si y sólo si no hay imperfecciones en una longitud de 3 milímetros; simple, pero suficiente para un análisis de la distribución de X.

En general, sea que la variable aleatoria N denote el número de imperfecciones en x milímetros de alambre. Si el número promedio de imperfecciones es λ por milímetro, entonces N tiene una distribución de Poisson con media λx. Se supone que el alambre es más largo que el valor de x. Ahora bien,

$$P(X > x) = P(N = 0) = \frac{e^{-\lambda x}(\lambda x)^0}{0!} = e^{-\lambda x}$$

Por lo tanto,

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

es la función de distribución acumulada de X. Al derivar $F(x)$, el cálculo de la función de densidad de probabilidad de X da como resultado

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

La deducción de la distribución de X depende únicamente del supuesto de que las imperfecciones en el alambre siguen un proceso de Poisson. Asimismo, el punto inicial para medir X no es relevante debido a que la probabilidad del número de imperfecciones en un intervalo de un proceso de Poisson depende tan sólo de la longitud del intervalo, no de la localización. Para cualquier proceso de Poisson, se aplica el siguiente resultado general:

Definición

La variable aleatoria X que es igual a la distancia entre conteos sucesivos de un proceso de Poisson con media $\lambda > 0$ tiene una **distribución exponencial** con parámetro λ. La función de densidad de probabilidad de X es

$$f(x) = \lambda e^{-\lambda x}, \quad \text{para} \quad 0 \leq x < \infty \tag{5-14}$$

La distribución exponencial debe su nombre a la función exponencial de la función de densidad de probabilidad. En la figura 5-24 se presentan gráficas de la distribución exponencial para valores escogidos de λ. Para cualquier valor de λ, la distribución exponencial presenta un sesgo considerable. Los resultados siguientes se obtienen fácilmente y se dejan como ejercicio.

Si la variable aleatoria X tiene una distribución exponencial con parámetro λ, entonces

$$E(X) = \frac{1}{\lambda} \quad \text{y} \quad V(X) = \frac{1}{\lambda^2} \tag{5-15}$$

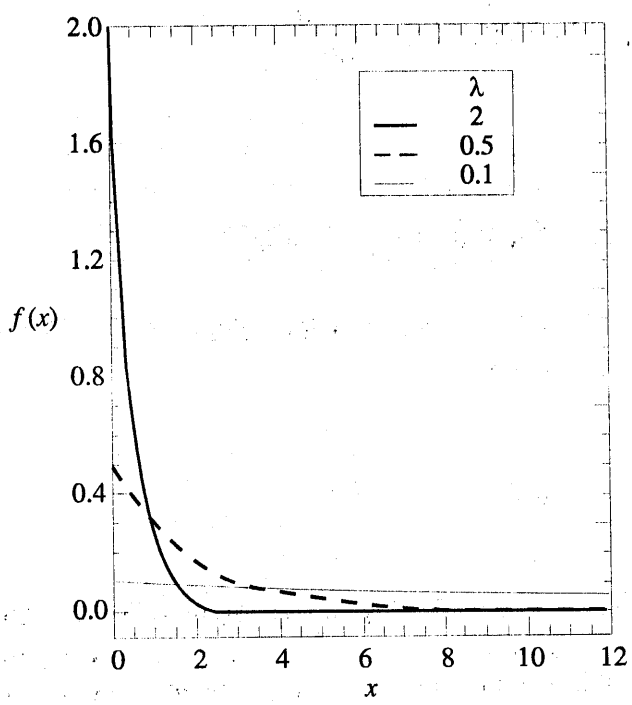


Figura 5-24 Función de densidad de probabilidad de una variable aleatoria exponencial para valores escogidos de λ.

Es importante **usar unidades consistentes** en el cálculo de probabilidades, medias y varianzas en las que intervengan variables aleatorias exponenciales. En los ejemplos siguientes se ilustra la conversión de unidades.

EJEMPLO 5-21

En una red de computadoras de una gran corporación, el acceso de usuarios al sistema puede modelarse como un proceso de Poisson con una media de 25 accesos por hora. ¿Cuál es la probabilidad de que no haya accesos en un intervalo de 6 minutos?

Sea que X denote el tiempo en horas desde el principio del intervalo hasta el primer acceso. Entonces, X tiene una distribución exponencial con $\lambda = 25$ accesos por hora. Nos interesa la probabilidad de que X exceda 6 minutos. Como λ está dada en accesos por hora, todas las unidades de tiempo se expresan en horas. Es decir, 6 minutos = 0.1 horas. La probabilidad pedida se indica con el área sombreada bajo la función de densidad de probabilidad de la figura 5-25. Por lo tanto,

$$P(X > 0.1) = \int_{0.1}^{\infty} 25e^{-25x}dx$$
$$= e^{-25(0.1)} = 0.082$$

Se llega exactamente al mismo resultado expresando el número promedio de accesos como 0.417 accesos por minuto y calculando la probabilidad de que el tiempo hasta el siguiente acceso exceda 6 minutos. ¡Inténtelo!

¿Cuál es la probabilidad de que el tiempo hasta el siguiente acceso esté entre 2 y 3 minutos? Después de convertir todas las unidades a horas,

$$P(0.033 < X < 0.05) = \int_{0.033}^{0.05} 25e^{-25x}dx$$
$$= -e^{-25x}\Big|_{0.033}^{0.05}$$
$$= 0.152$$

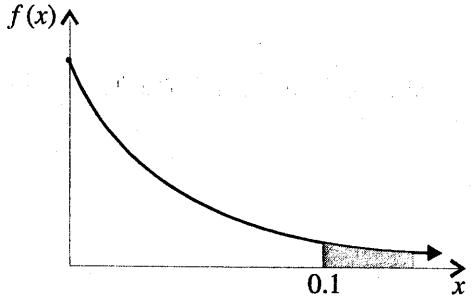


Figura 5-25 Probabilidad para la distribución exponencial del ejemplo 5-21.

Determine el intervalo de tiempo tal que la probabilidad de que no haya accesos en el intervalo sea 0.90. Se pide la longitud de tiempo x tal que $P(X > x) = 0.90$. Ahora bien,

$$P(X > x) = e^{-25x} = 0.90$$

Por lo tanto, después de sacar logaritmos en ambos miembros

$$x = 0.00421 \text{ horas} = 0.25 \text{ minutos}$$

Por otra parte, el tiempo promedio hasta el siguiente acceso es

$$\mu = 1/25 = 0.04 \text{ horas} = 2.4 \text{ minutos}$$

La desviación estándar del tiempo hasta el siguiente acceso es

$$\sigma = 1/25 \text{ horas} = 2.4 \text{ minutos}$$

En el ejemplo anterior, la probabilidad de que no haya accesos en un intervalo de 6 minutos es 0.082, independientemente del tiempo inicial del intervalo. En un proceso de Poisson se supone que los eventos ocurren de manera uniforme en todo el intervalo de observación; es decir, no hay amontonamiento de eventos. Si un proceso de Poisson es un modelo apropiado de los accesos, la probabilidad de que el primer acceso después de mediodía ocurra después de las 12:06 pm es la misma que la probabilidad de que el primer acceso después de las 3:00 pm ocurra a las 3:06 pm. Y si alguien tiene acceso al sistema a las 2:22 pm, la probabilidad de que el siguiente acceso ocurra después de las 2:28 pm sigue siendo 0.082.

El punto inicial para observar el sistema no importa. Sin embargo, si hay periodos de uso intenso durante el día, como justo después de las 8:00 am, seguido por un periodo de uso reducido, un proceso de Poisson no es un modelo apropiado para los accesos y la distribución no es adecuada para calcular probabilidades. Sería razonable modelar los periodos de uso intenso y reducido con un proceso de Poisson separado, empleando un valor mayor de λ durante los periodos de uso intensivo y un valor menor en caso contrario. Entonces, una distribución exponencial con el valor correspondiente de λ puede usarse para calcular las probabilidades de acceso para los periodos de uso intenso y reducido.

Propiedad de falta de memoria

Una propiedad aún más interesante de una variable aleatoria exponencial se relaciona con las probabilidades condicionales.

EJEMPLO 5-22

Sea que X denote el tiempo entre la detección de una partícula rara en un contador Geiger y suponga que X tiene una distribución exponencial con $\lambda = 1.4$ minutos. La probabilidad de que se detecte una partícula en un lapso de 30 segundos después de activar el contador es

$$P(X < 0.5 \text{ minutos}) = 1 - e^{-0.5/1.4} = 0.30$$

En este cálculo se ha hecho la conversión de todas las unidades a minutos. Ahora bien, suponga que se enciende el contador Geiger y que pasan 3 minutos sin que se detecte ninguna partícula. ¿Cuál es la probabilidad de detectar una partícula en los siguientes 30 segundos?

Como ya han transcurrido 3 minutos de espera, quizá se piense que uno "va a la segura", es decir, que la probabilidad de una detección positiva en los siguientes 30 segundos deberá ser mayor que 0.3. Sin embargo, para una distribución exponencial éste no es el caso. La probabilidad pedida puede expresarse como la probabilidad condicional $P(X < 3.5 \mid X > 3)$. Por la definición de probabilidad condicional,

$$P(X < 3.5 \mid X > 3) = P(3 < X < 3.5) / P(X > 3)$$

donde

$$\begin{aligned} P(3 < X < 3.5) &= F(3.5) - F(3) \\ &= \left[1 - e^{-3.5/1.4}\right] - \left[1 - e^{-3/1.4}\right] \\ &= 0.0035 \end{aligned}$$

y

$$P(X > 3) = e^{-3/1.4} = 0.117$$

Por lo tanto,

$$P(X < 3.5 \mid X > 3) = 0.035 / 0.117 = 0.30$$

Después de esperar 3 minutos sin una detección positiva, la probabilidad de una detección en los siguientes 30 segundos es igual que la probabilidad de una detección en los 30 primeros segundos después de encender el contador Geiger. El hecho de que se haya esperado 3 minutos sin una detección, no cambia la probabilidad de una detección en los 30 segundos siguientes.

En el ejemplo 5-22 se ilustra la **propiedad de falta de memoria** de una variable aleatoria exponencial. De hecho, la distribución exponencial es la única distribución continua con esta propiedad.

Propiedad de falta de memoria

Para una variable aleatoria exponencial X,

$$P(X < t_1 + t_2 \mid X > t_1) = P(X < t_2) \tag{5-16}$$

En la figura 5-26 se ilustra la propiedad de falta de memoria. El área de la región A dividida por el área total bajo la función de densidad de probabilidad ($A + B + C + D = 1$) es igual a $P(X < t_2)$. El área de la región C dividida por el área $C + D$ es igual a $P(X < t_1 + t_2 \mid X > t_1)$. La propiedad de falta de memoria es equivalente al hecho siguiente. La proporción del área total que está en A

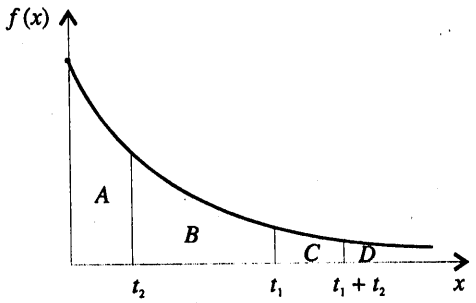


Figura 5-26 Propiedad de falta de memoria de una distribución exponencial.

es igual a la proporción del área en C y D que está en C. La verificación matemática de la propiedad de falta de memoria se deja como parte de los ejercicios para desarrollar el intelecto.

La propiedad de falta de memoria deja de ser tan sorprendente cuando se considera el desarrollo del proceso de Poisson. En ese desarrollo, se supuso que podía hacerse la partición de un intervalo en intervalos pequeños que fueran independientes. Estos subintervalos son similares a ensayos de Bernoulli independientes que comprenden un proceso binomial; conocer resultados previos no afecta las probabilidades de los eventos en subintervalos futuros. Una variable aleatoria exponencial es el análogo continuo de una variable aleatoria geométrica, y comparten una propiedad de falta de memoria similar.

La distribución exponencial se usa con frecuencia en estudios de confiabilidad como modelo para el tiempo hasta la falla de un dispositivo. Por ejemplo, la vida de un chip semiconductor podría modelarse como una variable aleatoria exponencial con una media de 40 000 horas. La propiedad de falta de memoria de la distribución exponencial implica que el dispositivo no se desgasta. Es decir, independientemente del tiempo que haya estado operando el dispositivo, la probabilidad de una falla en las 1 000 horas siguientes es igual que la probabilidad de una falla en las primeras 1 000 horas de operación. La vida L de un dispositivo con fallas causadas por imprevistos aleatorios podría modelarse adecuadamente como una variable aleatoria exponencial. Sin embargo, la vida L de un dispositivo que sufre desgaste mecánico lento, como el desgaste de un rodamiento, se modela mejor con una distribución tal que $P(L < t + \Delta t \mid L > t)$ se incremente con t. En la práctica, distribuciones como la de Weibull se usan con frecuencia para modelar el tiempo de falla de este tipo de dispositivo. La distribución de Weibull se presenta en la sección 5-11.

EJERCICIOS DE LA SECCIÓN 5-9

5-69. Suponga que X tiene una distribución exponencial con $\lambda = 2$. Determine lo siguiente:

a) $P(X \leq 0)$
b) $P(X \geq 2)$
c) $P(X \leq 1)$
d) $P(1 < X < 2)$
e) Encuentre el valor de x tal que $P(X < x) = 0.05$.

5-70. Suponga que X tiene una distribución exponencial con media igual a 10. Determine lo siguiente:

a) $P(X > 10)$
b) $P(X > 20)$
c) $P(X > 30)$
d) Encuentre el valor de x tal que $P(X < x) = 0.95$.

5-71. Suponga que los conteos registrados por un contador Geiger siguen un proceso de Poisson con un promedio de dos conteos por minuto.

a) ¿Cuál es la probabilidad de que no haya conteos en un intervalo de 30 segundos?

b) ¿Cuál es la probabilidad de que el primer conteo ocurra en menos de 10 segundos?

c) ¿Cuál es la probabilidad de que el primer conteo ocurra entre 1 y 2 minutos después de encender el contador?

5-72. Suponga que el acceso a una red de computadoras sigue un proceso de Poisson con un promedio de 3 conteos por minuto.

a) ¿Cuál es el tiempo promedio entre los conteos?

b) ¿Cuál es la desviación estándar del tiempo entre los conteos?

c) Determine x tal que la probabilidad de que ocurra al menos un conteo antes del tiempo x minutos sea 0.95.

5-73. El tiempo entre las llamadas telefónicas a una ferretería tiene una distribución exponencial con un tiempo promedio entre las llamadas de 15 minutos.

a) ¿Cuál es la probabilidad de que no haya llamadas en un intervalo de 30 minutos?

b) ¿Cuál es la probabilidad de que haya al menos una llamada en un intervalo de 10 minutos?

c) ¿Cuál es la probabilidad de que la primera llamada se realice dentro de los 5 y 10 minutos después de abrir?

d) Determine la longitud de un intervalo de tiempo tal que la probabilidad de que haya al menos una llamada en el intervalo sea 0.90.

5-74. La vida de los reguladores de voltaje para automóviles tiene una distribución exponencial con una vida media de seis años. Usted compra un automóvil con seis años de antigüedad que tiene el regulador de voltaje funcionando y planea conservarlo por seis años.

a) ¿Cuál es la probabilidad de que el regulador de voltaje falle antes de que venda el automóvil?

b) Si el regulador falla después de tres años de haber comprado el automóvil y lo reemplaza, ¿cuál es el tiempo promedio antes de la siguiente falla?

5-75. El tiempo entre la entrada de correos electrónicos en una computadora tiene una distribución exponencial con una media de dos horas.

a) ¿Cuál es la probabilidad de que no se reciba un correo electrónico durante un periodo de dos horas?

b) Si no se ha recibido un correo en las últimas cuatro horas, ¿cuál es la probabilidad de que no se reciba uno en las dos horas siguientes?

c) ¿Cuál es el tiempo esperado entre el quinto y el sexto correo electrónico?

5-76. El tiempo para que pase un taxi desocupado por un crucero muy transitado tiene una distribución exponencial con una media de 10 minutos.

a) ¿Cuál es la probabilidad de que tenga que esperar más de una hora por un taxi?

b) Suponga que una persona ha esperado ya una hora por un taxi; ¿cuál es la probabilidad de que pase uno en los próximos 10 minutos?

5-77. Continuación del ejercicio 5-76.

a) Determine x tal que la probabilidad de que una persona aguarde más de x minutos sea 0.10.

b) Determine x tal que la probabilidad de que una persona aguarde menos de x minutos sea 0.90.

c) Determine x tal que la probabilidad de que una persona aguarde menos de x minutos sea 0.50.

5-78. La distancia entre grietas grandes en una carretera sigue una distribución exponencial con una media de 5 millas.

a) ¿Cuál es la probabilidad de que no haya grietas grandes en un tramo de 10 millas de la carretera?

b) ¿Cuál es la probabilidad de que haya dos grietas grandes en un tramo de 10 millas de la carretera?

c) ¿Cuál es la desviación estándar de la distancia entre las grietas grandes?

5-79. Continuación del ejercicio 5-78.

a) ¿Cuál es la probabilidad de que la primera grieta grande se presente entre 12 y 15 millas después del inicio de la inspección?

b) ¿Cuál es la probabilidad de que no haya grietas grandes en dos tramos separados de 5 millas de la carretera?

c) Dado que no hay grietas grandes en las primeras 5 millas inspeccionadas, ¿cuál es la probabilidad de que no haya grietas grandes en las siguientes 10 millas inspeccionadas?

5-80. La vida de un ensamblaje mecánico en una prueba de vibración tiene una distribución exponencial con una media de 400 horas.

a) ¿Cuál es la probabilidad de que un ensamblaje sometido a prueba falle en menos de 100 horas?

b) ¿Cuál es la probabilidad de que un ensamblaje opere durante más de 500 horas antes de fallar?

c) Si un ensamblaje se ha sometido a prueba durante más de 400 horas sin fallar, ¿cuál es la probabilidad de una falla en las 100 horas siguientes?

5-81. Continuación del ejercicio 5-80.

a) Si se someten a prueba 10 ensamblajes, ¿cuál es la probabilidad de que al menos uno falle en menos de 100 horas? Suponga que los ensamblajes fallan independientemente.

b) Si se someten a prueba 10 ensamblajes, ¿cuál es la probabilidad de que todos hayan fallado en 800 horas? Suponga que los ensamblajes fallan independientemente.

5-82. Cuando una línea de autobuses reduce sus tarifas, un viaje en particular de la ciudad de Nueva York a Albany, Nueva York, se vuelve muy popular. Un autobús pequeño puede transportar cuatro pasajeros. El tiempo entre las llamadas telefónicas para comprar boletos tiene una distribución exponencial con una media de 30 minutos. Suponga que en cada llamada se compra un boleto. ¿Cuál es la probabilidad de que el autobús se llene en menos de 3 horas después de haberse iniciado la reducción de las tarifas?

5-83. El tiempo entre el aterrizaje de avionetas en un aeropuerto local tiene una distribución exponencial con una media de una hora. ¿Cuál es la probabilidad de que aterricen más de tres avionetas en una hora?

5-84. Continuación del ejercicio 5-83.

a) Si se eligen 30 intervalos separados de una hora, ¿cuál es la probabilidad de que ninguno de los intervalos contenga más de tres aterrizajes?

b) Determine la longitud del intervalo de tiempo (en horas) tal que la probabilidad de que no haya ningún aterrizaje durante el intervalo sea 0.10.

5-85. El tiempo entre las llamadas telefónicas a una oficina corporativa tiene una distribución exponencial con una media de 10 minutos.

a) ¿Cuál es la probabilidad de que haya más de tres llamadas en media hora?

b) ¿Cuál es la probabilidad de que no haya llamadas en media hora?

c) Determine x tal que la probabilidad de que no haya llamadas en x horas sea 0.01.

5-86. Continuación del ejercicio 5-85.

a) ¿Cuál es la probabilidad de que no haya llamadas en un intervalo de dos horas?

b) Si se seleccionan cuatro intervalos no traslapados de media hora, ¿cuál es la probabilidad de que en ninguno de estos intervalos haya llamadas?

c) Explique la relación entre los resultados de los incisos a) y b).

5-87. Si la variable aleatoria X tiene una distribución exponencial con media θ, determine lo siguiente:

a) $P(X > \theta)$

b) $P(X > 2\theta)$

c) $P(X > 3\theta)$

d) ¿En qué forma dependen los resultados de θ?

5-88. Suponga que los defectos en una cinta magnética siguen una distribución de Poisson con

una media de 0.2 defectos por metro. Sea que X denote la distancia entre dos defectos sucesivos.

a) ¿Cuál es la media de X?

b) ¿Cuál es la probabilidad de que no haya defectos en 10 metros consecutivos de cinta?

c) ¿La respuesta del inciso b) cambia si los 10 metros no son consecutivos?

d) ¿Cuántos metros de cinta es necesario inspeccionar para que la probabilidad de encontrar al menos un defecto sea de 90%?

5-89. Continuación del ejercicio 5-88. (Preguntas más difíciles.)

a) ¿Cuál es la probabilidad de que la primera ocasión en que la distancia entre dos defectos exceda 8 metros ocurra en el quinto defecto?

b) ¿Cuál es el número promedio de defectos antes de que la distancia entre dos defectos exceda 8 metros?

5-90. Deduzca la fórmula de la media y la varianza de una variable aleatoria exponencial.

5-10 DISTRIBUCIONES DE ERLANG Y GAMMA

5-10.1 Distribución de Erlang

Una variable aleatoria exponencial describe la longitud hasta que se obtiene el primer conteo en un proceso de Poisson. Una generalización de la distribución exponencial es la longitud hasta que ocurran r conteos en un proceso de Poisson. La variable aleatoria que es igual a la longitud del intervalo hasta que ocurran r conteos en un proceso de Poisson es una **variable aleatoria de Erlang**.

EJEMPLO 5-23

Las fallas de las unidades de procesamiento central (CPU) de los sistemas de computadoras grandes se modelan con frecuencia como un proceso de Poisson. Por lo general, las fallas no son causadas por el desgaste de componentes, sino por fallas de carácter más aleatorio en el gran número de circuitos semiconductores de las unidades. Suponga que las unidades que fallan se reparan de inmediato y que el número promedio de fallas por hora es 0.0001. Sea que X denote el tiempo hasta que ocurren cuatro fallas en un sistema. Determine la probabilidad de que X exceda 40 000 horas.

Sea la variable aleatoria N que denote el número de fallas en 40 000 horas de operación. El tiempo hasta que ocurran cuatro fallas excede 40 000 horas si y sólo si el número de fallas en 40 000 horas es tres o menos. Por lo tanto,

$$P(X > 40\,000) = P(N \le 3)$$

El supuesto de que las fallas siguen un proceso de Poisson implica que N tiene una distribución de Poisson con

$$E(N) = 40\,000(0.0001) = 4 \text{ fallas por } 40\,000 \text{ horas}$$

Por lo tanto,

$$P(X > 40\,000) = P(N \le 3) = \sum_{k=0}^{3} \frac{e^{-4}4^k}{k!} = 0.433$$

La función de distribución acumulada de una variable aleatoria de Erlang X puede obtenerse calculando $P(X \leq x) = 1 - P(X > x)$ y la probabilidad $P(X > x)$ puede determinarse como en el ejemplo anterior. Entonces, la función de densidad de probabilidad de X puede obtenerse derivando la función de distribución acumulada y realizando una gran cantidad de simplificaciones algebraicas. Los detalles se dejan como ejercicio al final de este capítulo. En general, puede llegarse al resultado siguiente:

Definición

La variable aleatoria X que es igual a la longitud del intervalo hasta que ocurren r fallas en un proceso de Poisson con media $\lambda > 0$ tiene una **distribución de Erlang** con parámetros λ y r. La función de densidad de probabilidad de X es

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!} \quad \text{para } x > 0 \text{ y } r = 1, 2, \ldots \tag{5-17}$$

En la figura 5-27 se muestran gráficas de la función de densidad de probabilidad de Erlang para varios valores de r y λ. Evidentemente, una variable aleatoria con $r = \lambda$ es una variable aleatoria exponencial. Es común determinar probabilidades en las que intervienen variables aleatorias de Erlang calculando una sumatoria de variables aleatorias de Poisson, como en el ejemplo 5-23. La función de densidad de probabilidad de una variable aleatoria de Erlang puede usarse para

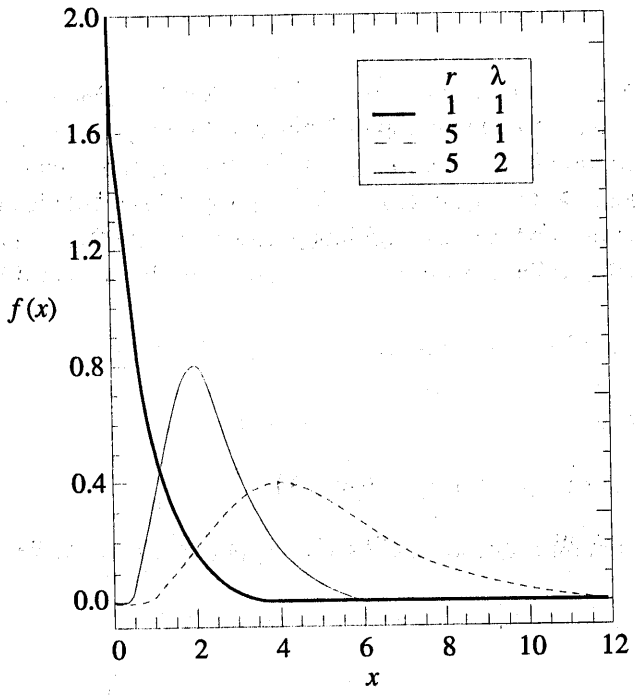


Figura 5-27 Funciones de densidad de probabilidad de Erlang para valores escogidos de r y λ.

determinar probabilidades; sin embargo, con frecuencia es necesario integrar por partes. Como en el caso de la distribución exponencial, debe tenerse cuidado en definir la variable aleatoria y el parámetro en unidades consistentes.

EJEMPLO 5-24

Una aproximación alternativa para calcular la probabilidad requerida en el ejemplo 5-23 es por integración de la función de densidad de probabilidad de X. Esto es

$$P(X > 40\,000) = \int_{40\,000}^{\infty} f(x)\,dx$$

$$= \int_{40\,000}^{\infty} \frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!}\,dx$$

donde $r = 4$ y $\lambda = 0.0001$. Se puede utilizar integración por partes para verificar los resultados obtenidos en el ejemplo 5-23.

Una variable aleatoria de Erlang puede considerarse como el análogo continuo de una variable aleatoria binomial negativa. Una variable aleatoria binomial negativa puede expresarse como la suma de r variables aleatorias geométricas. De manera similar, una variable aleatoria de Erlang puede representarse como la suma de r variables aleatorias exponenciales. Aplicando esta conclusión, puede demostrarse con facilidad el siguiente resultado:

Si X es una variable aleatoria de Erlang con parámetros λ y r, entonces la media y la varianza de X son

$$\mu = E(X) = r/\lambda \quad \text{y} \quad \sigma^2 = V(X) = r/\lambda^2 \tag{5-18}$$

5-10.2 Distribución gamma

La distribución de Erlang es un caso especial de la **distribución gamma**. Si el parámetro r de una variable aleatoria de Erlang no es un entero, pero $r > 0$, entonces la variable aleatoria tiene una distribución gamma. Sin embargo, en la función de densidad de Erlang, el parámetro r aparece como r factorial. Por lo tanto, para definir una variable aleatoria gamma, se requiere una generalización de la función factorial.

Definición

La **función gamma** es

$$\Gamma(r)=\int_{x=0}^{\infty} x^{r-1}e^{-x}dx, \quad \text{para} \quad r>0 \tag{5-19}$$

Puede demostrarse que la integral en la definición de $\Gamma(r)$ es finita. Además, al integrar por partes puede demostrarse que

$$\Gamma(r)=(r-1)\Gamma(r-1)$$

Este resultado se deja como ejercicio. Por lo tanto, si r es un entero positivo (como en la distribución de Erlang)

$$\Gamma(r)=(r-1)!$$

Asimismo, $\Gamma(1) = 0! = 1$ y puede demostrarse que $\Gamma(1/2) = \pi^{1/2}$. La función gamma puede interpretarse como una generalización para valores no enteros de r del término $(r - 1)!$, que se usa en la función de densidad de probabilidad de Erlang.

Ahora puede enunciarse la función de densidad de probabilidad gamma.

Definición

La variable aleatoria X con función de densidad de probabilidad

$$f(x)=\frac{\lambda^r x^{r-1}e^{-\lambda x}}{\Gamma(r)}, \quad \text{para} \quad x>0 \tag{5-20}$$

tiene una **distribución gamma** con parámetros $\lambda > 0$ y $r > 0$. Si r es un entero, entonces X tiene una distribución de Erlang.

En la figura 5-28 se muestran gráficas de la distribución gamma para varios valores de λ y r. Se deja como ejercicio demostrar que $f(x)$ satisface las propiedades de una función de densidad de probabilidad. Asimismo, puede llegarse al siguiente resultado:

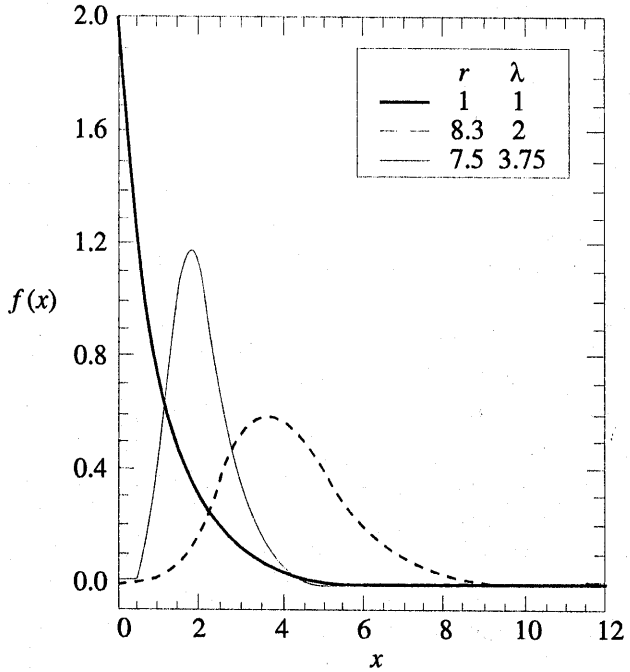


Figura 5-28 Funciones de densidad de probabilidad gamma para valores escogidos de λ y r.

Si X es una variable aleatoria gamma con parámetros λ y r, entonces la media y la varianza de X son

$$\mu = E(X) = r/\lambda \text{ y } \sigma^2 = V(X) = r/\lambda^2 \qquad (5\text{-}21)$$

Aun cuando la distribución gamma no se usa con frecuencia como modelo de un sistema físico, el caso especial de la distribución de Erlang es de suma utilidad para modelar experimentos aleatorios. Además, la **distribución ji-cuadrada** es un caso especial de la distribución gamma en la que $\lambda = 1/2$ y r es igual a los valores 1/2, 1, 3/2, 2, ... Esta distribución se usa ampliamente en la estimación de intervalos y en la prueba de hipótesis que se exponen en capítulos posteriores.

EJERCICIOS DE LA SECCIÓN 5-10

5-91. Las llamadas a un sistema telefónico siguen una distribución de Poisson con una media de cinco llamadas por minuto.

a) ¿Qué nombre se aplica a la distribución y a los valores de los parámetros del tiempo hasta la décima llamada?

b) ¿Cuál es el tiempo promedio hasta la décima llamada?

c) ¿Cuál es el tiempo promedio entre la novena y la décima llamada?

5-92. Continuación del ejercicio 5-91.

a) ¿Cuál es la probabilidad de que ocurran exactamente cuatro llamadas en un minuto?

b) Si se escogen 10 intervalos separados en un minuto, ¿cuál es la probabilidad de

que todos los intervalos contengan más de dos llamadas?

5-93. Se estudia la contaminación de una materia prima. Suponga que el número de partículas de contaminación por libra de material es una variable aleatoria de Poisson con una media de 0.01 partículas por libra.

a) ¿Cuál es el número esperado de libras necesarias de materia prima para obtener 15 partículas de contaminación?

b) ¿Cuál es la desviación estándar de las libras necesarias de materia prima para obtener 15 partículas de contaminación?

5-94. En un sistema de comunicación de datos, varios mensajes que llegan a un nodo se agrupan en un paquete antes de transmitirse en la red. Suponga que los mensajes que llegan al nodo siguen un proceso de Poisson con $\tau = 30$ mensajes por minuto. Se usan cinco mensajes para formar un paquete.

a) ¿Cuál es el tiempo promedio hasta que se forma un paquete, es decir, hasta que llegan cinco mensajes al nodo?

b) ¿Cuál es la desviación estándar del tiempo hasta que se forma un paquete?

c) ¿Cuál es la probabilidad de que se forme un paquete en menos de 10 segundos?

d) ¿Cuál es la probabilidad de que se forme un paquete en menos de 5 segundos?

5-95. Los errores causados por contaminación en discos ópticos ocurren a razón de un error cada 10^5 bits. Suponga que los errores siguen una distribución de Poisson.

a) ¿Cuál es el número promedio de bits hasta que ocurren cinco errores?

b) ¿Cuál es la desviación estándar del número de bits hasta que ocurren cinco errores?

c) El código de corrección de errores podría fallar si hay tres o más errores en 10^5 bits. ¿Cuál es la probabilidad de este evento?

5-96. El tiempo entre las llegadas de clientes a un cajero automático es una variable aleatoria exponencial con una media de 5 minutos.

a) ¿Cuál es la probabilidad de que lleguen más de tres clientes en 10 minutos?

b) ¿Cuál es la probabilidad de que el tiempo hasta que llegue el quinto cliente sea menor que 15 minutos?

5-97. El tiempo entre los problemas de procesamiento en una línea de producción, tiene una distribución exponencial con una media de 30 días.

a) ¿Cuál es el tiempo esperado hasta el cuarto problema?

b) ¿Cuál es la probabilidad de que el tiempo hasta el cuarto problema exceda 120 días?

5-98. Use las propiedades de la función gamma para evaluar lo siguiente:

a) $\Gamma(6)$

b) $\Gamma(5/2)$

c) $\Gamma(9/2)$

5-99. Use la integración por partes para demostrar que $\Gamma(r) = (r-1)\,\Gamma(r-1)$.

5-100. Demuestre que la integral de la función de densidad gamma $f(x; \lambda, r)$, es uno.

5-101. Use el resultado para la distribución gamma para determinar la media y la varianza de una distribución *ji*-cuadrada con $r = 7/2$.

5-11 DISTRIBUCIÓN DE WEIBULL

Como se mencionó anteriormente, la distribución de Weibull se usa con frecuencia para modelar el tiempo hasta que ocurre una falla en muchos sistemas físicos diferentes. Los parámetros de la distribución confieren una gran flexibilidad para modelar sistemas en los que el número de fallas aumenta con el tiempo (desgaste de rodamientos), disminuye con el tiempo (algunos semiconductores) o permanece constante con el tiempo (fallas causadas por interferencias externas al sistema).

Definición

La variable aleatoria X con función de densidad de probabilidad

$$f(x)=\frac{\beta}{\delta}\left(\frac{x}{\delta}\right)^{\beta-1} e^{-(x/\delta)^{\beta}}, \text{ para } x>0 \tag{5-22}$$

tiene una **distribución de Weibull** con parámetro de escala $\delta > 0$ y parámetro de forma $\beta > 0$.

La flexibilidad de la distribución de Weibull se ilustra en las gráficas de funciones de densidad de probabilidad escogidas de la figura 5-29. Por inspección de la función de densidad de probabilidad, se observa que cuando $\beta = 1$, la distribución de Weibull es idéntica a la distribución exponencial.

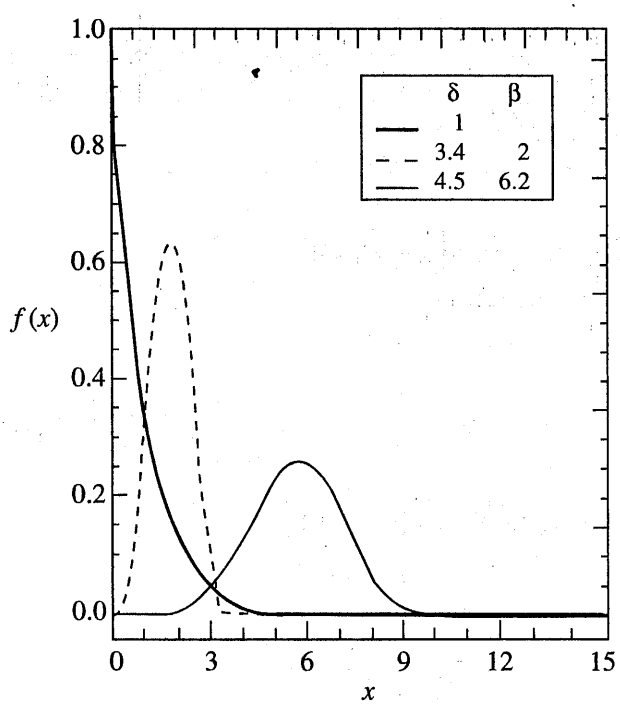


Figura 5-29 Funciones de densidad de probabilidad de Weibull para valores escogidos de δ y β.

Es común usar la función de distribución acumulada para calcular probabilidades. Puede llegarse al siguiente resultado:

Si X tiene una distribución de Weibull con parámetros δ y β, entonces la función de distribución acumulada de X es

$$F(x) = 1 - e^{-\left(\frac{x}{\delta}\right)^{\beta}} \tag{5-23}$$

También puede obtenerse el siguiente resultado:

Si X tiene una distribución de Weibull con parámetros δ y β, entonces la media y la varianza de x son

$$\mu = \delta\Gamma\left(1+\frac{1}{\beta}\right) \text{ y } \sigma^2 = \delta^2\Gamma\left(1+\frac{2}{\beta}\right) - \delta^2\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2 \tag{5-24}$$

EJEMPLO 5-25

El tiempo para que ocurra una falla (en horas) de un rodamiento en un eje mecánico se modela satisfactoriamente como una variable aleatoria de Weibull con $\beta = 1/2$ y $\delta = 5\,000$ horas. Determine el tiempo promedio para que ocurra una falla.

De la expresión para la media,

$$\begin{aligned} E(X) &= 5\,000\Gamma[1+(1/0.5)] \\ &= 5\,000\Gamma[3] \\ &= 5\,000\times 2! = 10\,000 \text{ horas} \end{aligned}$$

Determine la probabilidad de que un rodamiento dure por lo menos 6 000 horas. Ahora bien,

$$\begin{aligned} P(X > 6\,000) &= 1 - F(6\,000) \\ &= e - \left(\frac{6\,000}{5\,000}\right)^{1/2} \\ &= e^{-1.2} = 0.301 \end{aligned}$$

En consecuencia, 30.1% de todos los rodamientos duran al menos 6 000 horas.

EJERCICIOS DE LA SECCIÓN 5-11

5-102. Suponga que X tiene una distribución de Weibull con $\beta = 0.2$ y $\delta = 100$ horas. Determine la media y la varianza de X.

5-103. Suponga que X tiene una distribución de Weibull con parámetros $\beta = 0.2$ y $\delta = 100$ horas. Determine lo siguiente:
a) $P(X < 10\ 000)$
b) $P(X > 5\ 000)$

5-104. Suponga que la vida de un rodamiento de rodillos sigue una distribución de Weibull con parámetros $\beta = 2$ y $\delta = 10\ 000$ horas.
a) Determine la probabilidad de que un rodamiento dure al menos 8 000 horas.
b) Determine el tiempo promedio hasta la falla de un rodamiento.
c) Si están en uso 10 rodamientos y las fallas ocurren de manera independiente, ¿cuál es la probabilidad de que los 10 rodamientos duren al menos 8 000 horas?

5-105. Considere que la vida de un disco magnético empacado, expuesto a gases corrosivos, tiene una distribución de Weibull con $\beta = 0.5$ y la vida media es 600 horas.
a) Determine la probabilidad de que los discos empacados duren al menos 500 horas.
b) Determine la probabilidad de que los discos empacados fallen antes de 400 horas.

5-106. La vida de una bomba de recirculación sigue una distribución de Weibull con parámetros $\beta = 2$ y $\delta = 700$ horas.
a) Determine la vida media de una bomba.
b) Determine la varianza de la vida de una bomba.
c) ¿Cuál es la probabilidad de que una bomba dure más que su media?

5-107. Si X es una variable aleatoria con distribución de Weibull, con $\beta = 1$ y $\delta = 1\ 000$, ¿cuál es el otro nombre de la distribución de X y cuál es la media de X?

Ejercicios complementarios

5-108. Suponga que $f(x) = 0.5x - 1$ para $2 < x < 4$. Determine lo siguiente:
a) $P(X < 2)$
b) $P(X > 3)$
c) $P(2.5 < X < 3.5)$

5-109. Encuentre la función de distribución acumulada de la variable aleatoria del ejercicio 5-108.

5-110. Determine la media y la varianza de la variable aleatoria del ejercicio 5-108.

5-111. El tiempo entre llamadas telefónicas tiene una distribución exponencial con un tiempo promedio entre las llamadas de 10 minutos.
a) ¿Cuál es la probabilidad de que el tiempo hasta la primera llamada sea menor que 5 minutos?
b) ¿Cuál es la probabilidad de que el tiempo de la segunda a la tercera llamada esté entre 5 y 15 minutos?
c) Determine la longitud del intervalo de tiempo tal que la probabilidad de al menos una llamada en el intervalo, sea 0.90.

5-112. Continuación del ejercicio 5-111.
a) Si no ha habido ninguna llamada en 10 minutos, ¿cuál es la probabilidad de que el tiempo hasta la siguiente llamada sea menor que 5 minutos?
b) ¿Cuál es la probabilidad de que no haya llamadas en los intervalos de 10:00 a 10:05, de 11:30 a 11:35 y de 2:00 a 2:05?

5-113. Continuación del ejercicio 5-111.
a) ¿Cuál es la probabilidad de que haya exactamente cuatro llamadas en un minuto?
b) ¿Cuál es el tiempo promedio hasta la quinta llamada?

5-114. La CPU de una computadora personal tiene un periodo de vida con una distribución exponencial con una vida media de seis años. Usted ha sido dueño de esta CPU por tres años. ¿Cuál es la probabilidad de que la CPU falle en los próximos tres años?

5-115. Continuación del ejercicio 5-114. Suponga que la compañía para la que usted trabaja ha tenido 10 CPU por tres años, y suponga que las CPU fallan de manera independiente ¿Cuál es la probabilidad de que al menos una de ellas falle en los próximos tres años?

5-116. Se identifican fibras de asbesto en una muestra de polvo con un microscopio electrónico, después de preparar la muestra. Suponga que el número de fibras es una variable aleatoria de Poisson y que el número promedio de fibras por centímetro cuadrado de superficie de polvo es 100. Se analiza una muestra de 800 centímetros cuadrados de polvo. Suponga que una celda de una malla particular bajo el microscopio representa 1/160 000 de la muestra.

a) ¿Cuál es la probabilidad de que se localice al menos una fibra en la celda de la malla?

b) ¿Cuál es la media del número de celdas de la malla que es necesario observar para localizar 10 que contengan fibras?

c) ¿Cuál es la desviación estándar del número de celdas de la malla que es necesario observar para localizar 10 que contengan fibras?

5-117. Sin un sistema de riego automático, la altura de las plantas dos semanas después de germinar tiene una distribución normal con una media de 2.5 centímetros y una desviación estándar de 0.5 centímetros.

a) ¿Cuál es la probabilidad de que la altura de una planta sea mayor que 3.5 centímetros?

b) ¿Cuál es la probabilidad de que la altura de una planta esté entre 2.0 y 3.0 centímetros?

c) ¿Qué altura excede 90% de las plantas?

5-118. Continuación del ejercicio 5-117. Con un sistema de riego automático, una planta crece a una altura de 3.5 centímetros dos semanas después de germinar.

a) ¿Cuál es la probabilidad de obtener una planta de esta altura o mayor a partir de la distribución de las alturas del ejercicio 5-117?

b) ¿Considera usted que el sistema de riego automático incrementa la altura de las plantas a la segunda semana después de la germinación?

5-119. El espesor de un recubrimiento laminado para superficies de madera, tiene una distribución normal con una media de 5 milímetros y una desviación estándar de 0.2 milímetros.

a) ¿Cuál es la probabilidad de que el espesor de un recubrimiento sea mayor que 5.5 milímetros?

b) Si las especificaciones requieren que el espesor esté entre 4.5 y 5.5 milímetros, ¿qué proporción de los recubrimientos no cumplen con las especificaciones?

c) ¿Abajo de qué valor está el espesor del recubrimiento de 90% de las muestras?

5-120. El diámetro de los puntos producidos por una impresora matricial tiene una distribución normal con un diámetro promedio de 0.002 pulgadas. Suponga que las especificaciones requieren que el diámetro de los puntos esté entre 0.0014 y 0.0026 pulgadas. Si la probabilidad de que un punto cumpla con las especificaciones debe ser 0.9973, ¿qué desviación estándar se necesita?

5-121. Continuación del ejercicio 5-120. Suponga que la desviación estándar del tamaño de los puntos es 0.0004 pulgadas. Si la probabilidad de que un punto cumpla con las especificaciones debe ser 0.9973, ¿qué especificaciones es necesario establecer? Suponga que las especificaciones van a escogerse simétricamente alrededor de la media de 0.002.

5-122. La vida de un láser de semiconductores con una alimentación de energía constante tiene una distribución normal con una media de 7 000 horas y una desviación estándar de 600 horas.

a) ¿Cuál es la probabilidad de que un láser falle antes de 5 800 horas?

b) ¿Cuál es la vida en horas que excede 90% de los láser?

5-123. Continuación del ejercicio 5-122. ¿Qué valor deberá tener la vida media para que 99% de los láser excedan 10 000 horas antes de fallar?

5-124. Continuación del ejercicio 5-122. Un producto contiene tres láser, y el producto falla si cualquiera de ellos falla. Suponga que los

láser fallan de manera independiente. ¿Qué valor deberá tener la vida media para que 99% de los productos excedan 10 000 horas antes de fallar?

5-125. Una pulgada cuadrada de alfombra contiene 50 fibras del tejido. La probabilidad de una fibra dañada es 0.0001. Suponga que las fibras dañadas ocurren de manera independiente.

a) Aproxime la probabilidad de una fibra dañada en una yarda cuadrada de alfombra.

b) Aproxime la probabilidad de cuatro o más fibras dañadas en una yarda cuadrada de alfombra.

5-126. Una aerolínea hace 200 reservaciones para un vuelo con capacidad para 185 pasajeros. La probabilidad de que un pasajero llegue para el vuelo es 0.9 y se supone que los pasajeros son independientes.

a) Aproxime la probabilidad de que haya asientos para todos los pasajeros que lleguen.

b) Aproxime la probabilidad de que haya asientos vacíos.

c) Aproxime el número de reservaciones que debería hacer la aerolínea para que la probabilidad de que haya asientos para todos los pasajeros que lleguen sea 0.95. [*Sugerencia:* Pruébense valores sucesivos para el número de reservaciones.]

EJERCICIOS PARA DESARROLLAR EL INTELECTO

5-127. Los incisos de este ejercicio llevan a la función de densidad de probabilidad de una variable aleatoria de Erlang X con parámetros λ y r, $f(x) = \lambda^r x^{r-1} e^{-\lambda x}/(r-1)!$, $x > 0$, $r = 1, 2, \ldots$.

a) Use la distribución de Poisson para expresar $P(X > x)$.

b) Use el resultado del inciso a) para determinar la función de distribución acumulada de X.

c) Derive la función de distribución acumulada del inciso b) y simplifique para obtener la función de densidad de probabilidad de X.

5-128. Un juego de rodamientos se compone de 10 rodamientos. Suponga que los diámetros de los rodamientos son independientes y que tienen una distribución normal con una media de 1.5 milímetros y una desviación estándar de 0.025 milímetros. ¿Cuál es la probabilidad de que el rodamiento con el diámetro más grande en el ensamble exceda 1.6 milímetros?

5-129. Sea que la variable aleatoria X denote la medición de un producto manufacturado. Suponga que el valor especificado para la medición es m. Por ejemplo, X denotaría una longitud dimensional, y el valor especificado podría ser 10 milímetros. La **pérdida de calidad** del proceso para fabricar el producto se define como el valor esperado de $\$k(X - m)^2$, donde k es una constante que relaciona una desviación del valor especificado con una pérdida cuantificada en unidades monetarias ($).

a) Suponga que X es una variable aleatoria continua con $E(X) = m$ y $V(X) = \sigma^2$. ¿Cuál es la pérdida de calidad del proceso?

b) Suponga que X es una variable aleatoria continua con $E(X) = \mu$ y $V(X) = \sigma^2$. ¿Cuál es la pérdida de calidad del proceso?

EJERCICIOS PARA DESARROLLAR EL INTELECTO

5-130. La vida de un amplificador electrónico se modela como una variable aleatoria exponencial. Si 10% de los amplificadores tienen una vida media de 20 000 horas y el resto tiene una vida media de 50 000 horas, ¿qué proporción de los amplificadores falla antes de 60 000 horas?

5-131. **Propiedad de falta de memoria.** Demuestre que para una variable aleatoria exponencial X,

$$P(X < t_1 + t_2 \mid X > t_1) = P(X < t_2)$$

5-132. Se dice que un proceso es de **calidad seis sigma** si la media del proceso está al menos a seis desviaciones estándar de la especificación más próxima. Suponga una medición con una distribución normal.

a) Si la media de un proceso se ubica en el centro entre la especificación superior y la inferior a una distancia de seis desviaciones estándar entre sí, ¿cuál es la probabilidad de que un producto no cumpla con las especificaciones? Usando el resultado de que 0.000001 es igual a una parte por millón, exprese la respuesta en partes por millón.

b) Debido a que es difícil mantener la media de un proceso centrada entre las especificaciones, con frecuencia la probabilidad de que un producto no cumpla con las especificaciones se calcula después de Suponga que el proceso se corre. Si la media del proceso colocada como se indica en el inciso a) cambia 1.5 desviaciones estándar hacia arriba, ¿cuál es la probabilidad de que un producto no cumpla con las especificaciones? Exprese la respuesta en partes por millón.

6 Distribución de probabilidad conjunta

PLAN GENERAL DEL CAPÍTULO

6-1 DOS VARIABLES ALEATORIAS DISCRETAS
- 6-1.1 Distribuciones de probabilidad conjunta
- 6-1.2 Distribuciones de probabilidad marginal
- 6-1.3 Distribuciones de probabilidad condicional
- 6-1.4 Independencia

6-2 VARIABLES ALEATORIAS DISCRETAS MÚLTIPLES
- 6-2.1 Distribuciones de probabilidad conjunta
- 6-2.2 Distribución de probabilidad multinomial

6-3 DOS VARIABLES ALEATORIAS CONTINUAS
- 6-3.1 Distribuciones de probabilidad conjunta
- 6-3.2 Distribuciones de probabilidad marginal
- 6-3.3 Distribuciones de probabilidad condicional
- 6-3.4 Independencia

6-4 VARIABLES ALEATORIAS CONTINUAS MÚLTIPLES

6-5 COVARIANZA Y CORRELACIÓN

6-6 DISTRIBUCIÓN NORMAL DE DOS VARIABLES

6-7 COMBINACIONES LINEALES DE VARIABLES ALEATORIAS

6-8 DESIGUALDAD DE CHEBYSHEV

En los capítulos 4 y 5 se estudiaron las distribuciones de probabilidad para una sola variable aleatoria. Sin embargo, con frecuencia resulta conveniente tener definida más de una variable aleatoria en un experimento aleatorio. Por ejemplo, en la clasificación de señales transmitidas o recibidas, la calidad de cada señal puede clasificarse como alta, media o baja. La variable

aleatoria X podría definirse como el número de señales recibidas de alta calidad y la variable Y como el número de señales recibidas de baja calidad. En otro ejemplo, la variable aleatoria continua X podría denotar la longitud de una dimensión de una pieza moldeada por inyección, mientras que la variable aleatoria continua Y podría denotar la longitud de otra dimensión. Podrían ser de interés las probabilidades que pueden expresarse en términos de X y Y. Por ejemplo, si las especificaciones de X y Y son (2.95 a 3.05) y (7.60 a 7.80) milímetros, respectivamente, entonces podría ser de interés la probabilidad de que una pieza satisfaga ambas especificaciones; es decir, $P(2.95 < X < 3.05 \text{ y } 7.60 < Y < 7.80)$.

En general, si X y Y son dos variables aleatorias, a la distribución de probabilidad que define su comportamiento simultáneo se le llama **distribución de probabilidad conjunta**. En el presente capítulo se investigan algunas propiedades importantes de estas distribuciones conjuntas.

6-1 DOS VARIABLES ALEATORIAS DISCRETAS

6-1.1 Distribuciones de probabilidad conjunta

A fin de simplificar la exposición, se empieza considerando experimentos aleatorios en los que sólo se estudian dos variables aleatorias. En secciones posteriores se generaliza la presentación de la distribución de probabilidad conjunta para el caso de más de dos variables aleatorias.

EJEMPLO 6-1

En el desarrollo de un nuevo receptor para la transmisión de información digital, cada bit recibido se clasifica como aceptable, dudoso o inaceptable, dependiendo de la calidad de la señal recibida, con probabilidades 0.9, 0.08 y 0.02, respectivamente. Suponga que la clasificación de cada bit es independiente.

En los primeros cuatro bits transmitidos, sea que X denote el número de bits aceptables y Y el número de bits dudosos. Entonces, la distribución de X es binomial con $n = 4$ y $p = 0.9$, y la distribución de Y es binomial con $n = 4$ y $p = 0.08$. Sin embargo, puesto que sólo se están clasificando cuatro bits, los valores posibles de X y Y se restringen a los puntos indicados en la gráfica de la figura 6-1. Aun cuando los valores posibles de X son 0, 1, 2, 3 o 4, si $y = 3$, entonces $x = 0$ o 1. Al especificar la probabilidad de cada uno de los puntos en la figura 6-1, se especifica la distribución de probabilidad conjunta de X y Y. Como en el caso de una sola variable aleatoria, el rango de las variables aleatorias (X, Y) se define como el conjunto de puntos (x, y) en el espacio bidimensional para los que la probabilidad de que $X = x$ y $Y = y$ es positiva.

Si X y Y son variables aleatorias discretas, la distribución de probabilidad conjunta de X y Y es una descripción del conjunto de puntos (x, y) en el rango (X, Y) junto con la probabilidad de cada punto. A la distribución de probabilidad conjunta de dos variables aleatorias en ocasiones se le llama **distribución de probabilidad de dos variables** o **distribución bivariable**, de las variables aleatorias. Una forma de describir la distribución de probabilidad conjunta de dos variables aleatorias discretas es utilizando una función de masa de probabilidad conjunta. Asimismo, $P(X = x \text{ y } Y = y)$ acostumbra escribirse $P(X = x, Y = y)$.

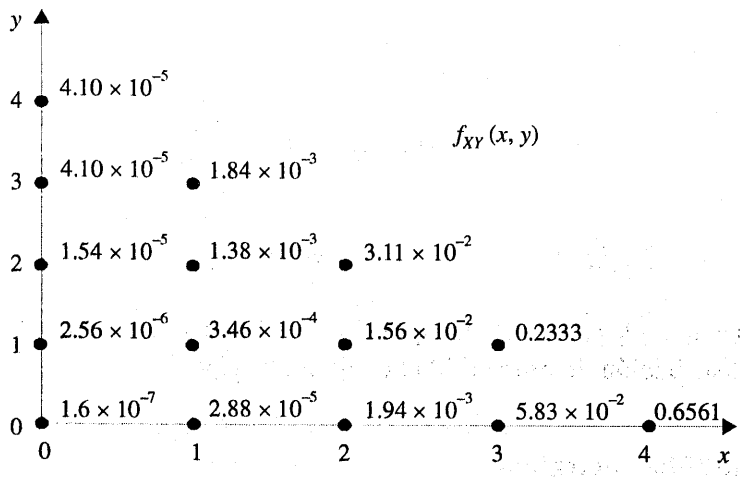


Figura 6-1 Distribución de probabilidad conjunta de X y Y en el ejemplo 6-1.

Definición

La **función de masa de probabilidad conjunta** de las variables aleatorias discretas X y Y, denotada como $f_{XY}(x, y)$, satisface

1) $f_{XY}(x, y) \geq 0$

2) $\sum_x \sum_y f_{XY}(x, y) = 1$

3) $f_{XY}(x, y) = P(X = x, Y = y)$ (6-1)

Los subíndices se usan para indicar las variables aleatorias en la distribución de probabilidad de dos variables. Así como en el caso de la función de masa de probabilidad de una sola variable aleatoria se supone que X es cero en todos los valores fuera del rango de X, se da por hecho que la función de masa de probabilidad conjunta de X y Y es cero en los valores para los que no está especificada una probabilidad.

EJEMPLO 6-2

La probabilidad de cada punto de la figura 6-1 se determina como sigue. Por ejemplo, $P(X = 2, Y = 1)$ es la probabilidad de que exactamente dos bits aceptables y exactamente un bit dudoso se reciban entre los cuatro bits transferidos. Sea que A, S y U denoten bits aceptables, dudosos e inaceptables, respectivamente. Por el supuesto de independencia,

$$P(AASU) = 0.9(0.9)(0.08)(0.02) = 0.0013$$

En el apéndice I se encuentra que el número de secuencias posibles que constan de dos A, una S y una U es

$$\frac{4!}{2!1!1!} = 12$$

Por lo tanto,

$$P(AASU) = 12(0.0013) = 0.0156$$

y

$$f_{XY}(2, 1) = P(X = 2, Y = 1) = 0.0156$$

Las probabilidades de todos los puntos de la figura 6-1 se indican en seguida del punto. La figura 6-1 describe la distribución de probabilidad conjunta de X y Y.

6-1.2 Distribuciones de probabilidad marginal

Si en un experimento aleatorio se define más de una variable aleatoria, es importante distinguir entre la distribución de probabilidad conjunta de X y Y y la distribución de probabilidad de cada variable individual. A la distribución de probabilidad individual de una variable aleatoria se le llama su **distribución de probabilidad marginal**. En el ejemplo 6-1 se mencionó que la distribución de probabilidad marginal de X es binomial con $n = 4$ y $p = 0.9$ y que la distribución de probabilidad marginal de Y es binomial con $n = 4$ y $p = 0.08$.

En general, la distribución de probabilidad marginal de X puede determinarse a partir de la distribución de probabilidad conjunta de X y otras variables aleatorias. Por ejemplo, para determinar $P(X = x)$, se suma $P(X = x, Y = y)$ en todos los puntos del rango de (X, Y) para los que $X = x$. Los subíndices de las funciones de masa de probabilidad sirven para distinguir las variables aleatorias.

EJEMPLO 6-3

La distribución de probabilidad conjunta de X y Y en la figura 6-1 puede usarse para encontrar la distribución de probabilidad marginal de X. Por ejemplo,

$$\begin{aligned} P(X = 3) &= P(X = 3, Y = 0) + P(X = 3, Y = 1) \\ &= 0.0583 + 0.2333 \\ &= 0.292 \end{aligned}$$

Como se esperaba, esta probabilidad coincide con el resultado obtenido a partir de la distribución de probabilidad binomial de X, es decir, $P(X = 3) = \binom{4}{3} 0.9^3 0.1^1 = 0.292$. La distribución de probabilidad marginal de X se encuentra sumando las probabilidades de cada columna, mientras que la distribución de probabilidad marginal de Y se encuentra sumando las probabilidades de cada renglón. En la figura 6-2 se muestran los resultados. Aun cuando la distribución de probabilidad marginal de X en este ejemplo puede determinarse directamente a partir de la descripción del experimento, en algunos problemas la distribución de probabilidad marginal sólo puede determinarse a partir de la distribución de probabilidad conjunta.

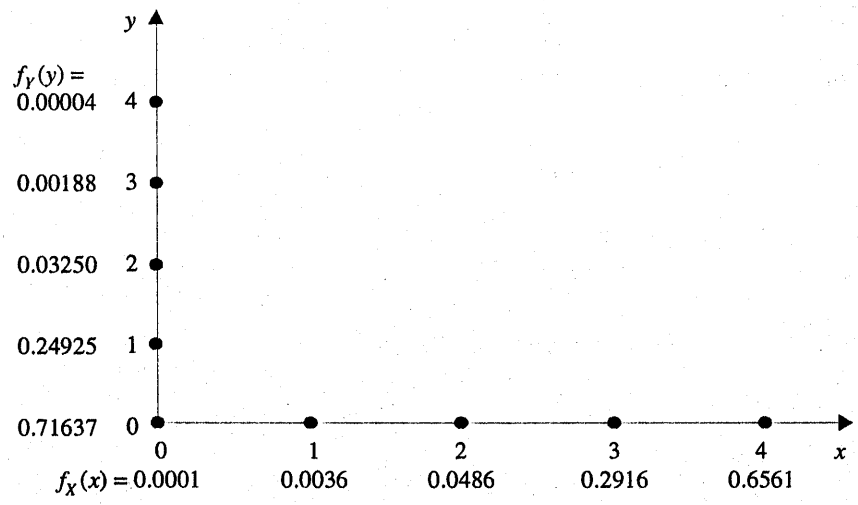


Figura 6-2 Distribuciones de probabilidad marginal de X y Y de la figura 6-1.

Definición

Si X y Y son variables aleatorias discretas con función de masa de probabilidad conjunta $f_{XY}(x, y)$, entonces las **funciones de masa de probabilidad marginal** de X y Y son

$$f_X(x) = P(X = x) = \sum_{R_x} f_{XY}(x, y)$$

$$f_Y(y) = P(Y = y) = \sum_{R_y} f_{XY}(x, y) \tag{6-2}$$

donde

R_x denota el conjunto de todos los puntos en el rango de (X, Y) para los que $X = x$ y

R_y denota el conjunto de todos los puntos en el rango de (X, Y) para los que $Y = y$

Dada una función de masa de probabilidad conjunta de las variables aleatorias X y Y, $E(X)$ y $V(X)$ pueden obtenerse directamente de la distribución de probabilidad conjunta de X y Y o bien calculando primero la distribución de probabilidad marginal de X y determinando después $E(X)$ y $V(X)$ aplicando el método usual. Lo anterior se ilustra en la siguiente ecuación:

Si la distribución de probabilidad marginal de X tiene la función de masa de probabilidad $f_X(x)$, entonces

$$E(X) = \mu_X = \sum_x x f_X(x) = \sum_x x\left(\sum_{R_x} f_{XY}(x, y)\right) = \sum_x \sum_{R_x} x f_{XY}(x, y)$$

$$= \sum_R x f_{XY}(x, y) \tag{6-3}$$

y

$$V(X) = \sigma_X^2 = \sum_x (x - \mu_X)^2 f_X(x) = \sum_x (x - \mu_X)^2 \sum_{R_x} f_{XY}(x, y)$$

$$= \sum_x \sum_{R_x} (x - \mu_X)^2 f_{XY}(x, y) = \sum_R (x, \mu_X)^2 f_{XY}(x, y)$$

donde

R_x denota el conjunto de todos los puntos en el rango de (X, Y) para los que $X = x$ y R denota el conjunto de todos los puntos en el rango de (X, Y)

EJEMPLO 6-4

En el ejemplo 6-1, $E(X)$ puede encontrarse como

$$\begin{aligned} E(X) &= 0[f_{XY}(0, 0) + f_{XY}(0, 1) + f_{XY}(0, 2) + f_{XY}(0, 3) + f_{XY}(0, 4)] \\ &\quad + 1[f_{XY}(1, 0) + f_{XY}(1, 1) + f_{XY}(1, 2) + f_{XY}(1, 3)] \\ &\quad + 2[f_{XY}(2, 0) + f_{XY}(2, 1) + f_{XY}(2, 2)] \\ &\quad + 3[f_{XY}(3, 0) + f_{XY}(3, 1)] \\ &\quad + 4[f_{XY}(4, 0)] \\ &= 0[0.0001] + 1[0.0036] + 2[0.0486] + 3\,[0.2916] + 4[0.6561] \\ &= 3.6 \end{aligned}$$

Alternativamente, debido a que la distribución de probabilidad marginal de X es binomial,

$$E(X) = np = 4(0.9) = 3.6$$

El cálculo, utilizando la distribución de probabilidad conjunta, puede usarse para determinar $E(X)$ incluso en los casos en que la distribución de probabilidad marginal de X no se conozca. Como ejercicio, el lector puede usar la distribución de probabilidad conjunta para comprobar que $E(Y) = 0.32$ en el ejemplo 6-1.

Asimismo,

$$V(X) = np(1 - p) = 4(0.9)(1 - 0.9) = 0.36$$

Compruebe que puede llegarse al mismo resultado a partir de la distribución de probabilidad conjunta de X y Y.

6-1.3 Distribuciones de probabilidad condicional

Cuando en un experimento aleatorio se definen dos variables aleatorias, la información que se tenga de una de ellas puede cambiar las probabilidades que se asocian con los valores de la otra. Recuérdese que en el ejemplo 6-1, X denota el número de bits aceptables y Y denota el número de bits dudosos recibidos por un receptor. Puesto que sólo se transmiten cuatro bits, si $X = 4$, entonces Y debe ser igual a 0. Utilizando la notación para las probabilidades condicionales del capítulo 3, se puede escribir este resultado como $P(Y = 0 \mid X = 4) = 1$. Si $X = 3$, entonces Y sólo puede ser 0 o 1. Por consiguiente, puede considerarse que las variables aleatorias X y Y son dependientes. Conocer el valor obtenido para X cambia las probabilidades asociadas con los valores de Y.

Recuérdese que la definición de probabilidad condicional para los eventos A y B es $P(A \mid B) = P(A \cap B)/P(A)$. Esta definición puede aplicarse al evento A definido como $X = x$ y al evento B definido como $Y = y$.

EJEMPLO 6-5

Para el ejemplo 6-1,

$$\begin{aligned} P(Y = 0 \mid X = 3) &= P(X = 3, Y = 0)/P(X = 3) \\ &= f_{XY}(3, 0)/f_X(3) \\ &= 0.05832/0.2916 = 0.200 \end{aligned}$$

La probabilidad de que $Y = 1$, dado que $X = 3$ es

$$\begin{aligned} P(Y = 1 \mid X = 3) &= P(X = 3, Y = 1)/P(X = 3) \\ &= f_{XY}(3, 1)/f_X(3) \\ &= 0.2333/0.2916 = 0.800 \end{aligned}$$

Dado que $X = 3$, los únicos valores posibles de Y son 0 y 1. Obsérvese que $P(Y = 0 \mid X = 3) + P(Y = 1 \mid X = 3) = 1$. Los valores 0 y 1 para Y junto con las probabilidades 0.200 y 0.800 definen la distribución de probabilidad condicional de Y, dado que $X = 3$.

El ejemplo 6-5 ilustra que las probabilidades condicionales de que $Y = y$, dado que $X = x$, puede considerarse como una nueva distribución de probabilidad. En la siguiente definición se generalizan estas ideas.

Definición

Dadas las variables aleatorias discretas X y Y con función de masa de probabilidad conjunta $f_{XY}(x, y)$, la **función de masa de probabilidad condicional** de Y, dado que $X = x$, es

$$f_{Y|x}(y) = f_{XY}(x, y)/f_X(x) \quad \text{para } f_X(x) > 0 \tag{6-4}$$

La función $f_{Y|x}(y)$ se usa para encontrar las probabilidades de los valores posibles de Y, dado que $X = x$. Es decir, es la función de masa de probabilidad de los valores posibles de Y, dado que $X = x$. En términos más precisos, sea que R_x denote el conjunto de todos los puntos en el rango de (X, Y) para los que $X = x$. La función de masa de probabilidad condicional proporciona las probabilidades condicionales de los valores de Y en el conjunto R_x.

Puesto que la función de masa de probabilidad condicional $f_{Y|x}(y)$ es una función de masa de probabilidad para toda y en R_x, se satisfacen las siguientes propiedades:

1) $f_{Y|x}(y) \geq 0$

2) $\sum_{R_x} f_{Y|x}(y) = 1$

3) $P(Y = y \mid X = x) = f_{Y|x}(y)$ (6-5)

EJEMPLO 6-6

Para la distribución de probabilidad conjunta de la figura 6-1, $f_{Y|x}(y)$ se encuentra dividiendo cada $f_{XY}(x, y)$ por $f_X(x)$. Aquí, $f_X(x)$ es simplemente la suma de las probabilidades de cada columna de la figura 6-1. La función $f_{Y|x}(y)$ se ilustra en la figura 6-3, donde la suma de cada columna es uno, debido a que es una distribución de probabilidad.

Las propiedades de las variables aleatorias pueden ampliarse a la distribución de probabilidad condicional de Y, dado que $X = x$.

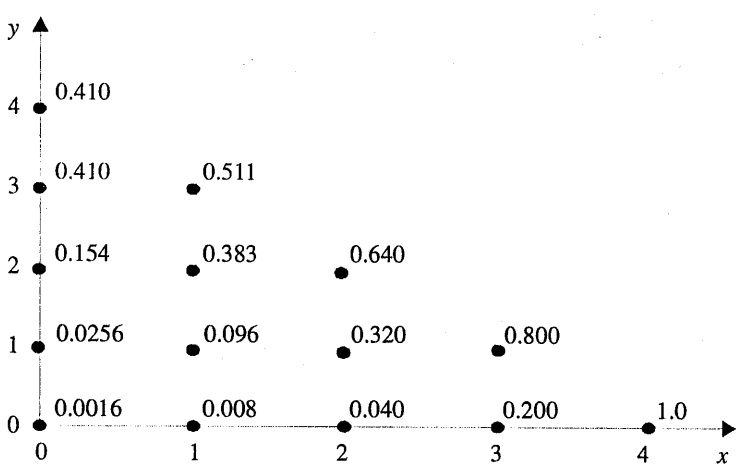


Figura 6-3 Distribuciones de probabilidad condicional de Y, dado que $X = x$, $f_{Y|x}(y)$, en el ejemplo 6-1.

Definición

Sea que R_x denote el conjunto de todos los puntos en el rango (X, Y) para los que $X = x$. La **media condicional** de Y, dado que $X = x$, denotada como $E(Y \mid x)$ o $\mu_{Y|x}$, es

$$E(Y|x) = \sum_{R_x} y f_{Y|x}(y) \tag{6-6a}$$

y la **varianza condicional** de Y, dado que $X = x$, denotada como $V(Y \mid x)$ o $\sigma^2_{Y|x}$, es

$$V(Y|x) = \sum_{R_x} (y - \mu_{Y|x})^2 f_{Y|x}(y) = \sum_{R_x} y^2 f_{Y|x}(y) - \mu^2_{Y|x} \tag{6-6b}$$

EJEMPLO 6-7

Para las variables aleatorias del ejemplo 6-1, la media condicional de Y, dado que $X = 2$, se obtiene a partir de la distribución condicional de la figura 6-3:

$$E(Y \mid 2) = \mu_{Y|2} = 0(0.040) + 1(0.320) + 2(0.640) = 1.6$$

La media condicional se interpreta como el número esperado de bits aceptables, dado que dos de los cuatro bits transmitidos son dudosos. La varianza condicional de Y, dado que $X = 2$, es

$$V(Y \mid 2) = (0 - \mu_{Y|2})^2(0.040) + (1 - \mu_{Y|2})^2 (0.320) + (2 - \mu_{Y|2})^2(0.640) = 0.32$$

6-1.4 Independencia

En algunos experimentos aleatorios, conocer los valores de X no cambia ninguna de las probabilidades asociadas con los valores de Y.

EJEMPLO 6-8

En una operación de moldeo de plástico, cada pieza se clasifica de acuerdo a si cumple o no con las especificaciones de color y longitud. La variable aleatoria X se define como

$$X = \begin{cases} 0 & \text{si la pieza cumple con las especificaciones de color} \\ 1 & \text{en caso contrario} \end{cases}$$

y la variable aleatoria Y se define como

$$Y = \begin{cases} 0 & \text{si la pieza cumple con las especificaciones de longitud} \\ 1 & \text{en caso contrario} \end{cases}$$

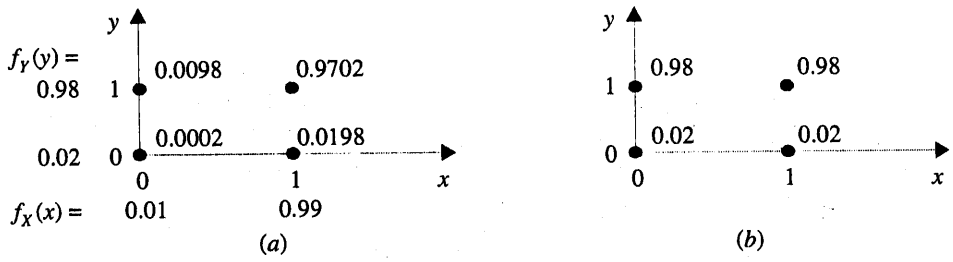


Figura 6-4 ***a***) Distribuciones de probabilidad conjunta y marginal de X y Y en el ejemplo 6-8. ***b***) Distribución de probabilidad condicional de Y, dado que $X = x$ en el ejemplo 6-8.

Suponga que la distribución de probabilidad conjunta de X y Y está definida por $f_{XY}(x, y)$ en la figura 6-4*a*. La distribución de probabilidad marginal de X y Y se muestra también en la figura 6-4*a*. Obsérvese que $f_{XY}(x, y) = f_X(x)f_Y(y)$. La función de masa de probabilidad condicional $f_{Y|x}(y)$ se ilustra en la figura 6-4*b*. Nótese que para toda x, $f_{Y|x}(y) = f_Y(y)$. Es decir, saber si la pieza cumple o no con las especificaciones de color, no cambia la probabilidad de que cumpla con las especificaciones de longitud.

Por analogía con un evento independiente, dos variables aleatorias se definen como **independientes** siempre que $f_{XY}(x, y) = f_X(x)f_Y(y)$ para toda x y y. Si dos variables aleatorias son independientes, entonces

$$f_{Y|x}(y) = f_{XY}(x, y)/f_X(x) = f_X(x)f_Y(y)/f_X(x) = f_Y(y)$$

Para las variables aleatorias discretas X y Y, si cualquiera de las propiedades siguientes es válida, entonces las demás también lo son, y X y Y son **independientes**.

1) $f_{XY}(x, y) = f_X(x)f_Y(y)$ para toda x y y

2) $f_{Y|x}(y) = f_Y(y)$ para toda x y y con $f_x(x) > 0$

3) $f_{X|y}(x) = f_X(x)$ para toda x y y con $f_Y(y) > 0$

4) $P(X = x, Y = y) = f_x(x) = f_y(y)$ para toda (x, y) en el rango de X y Y (6-7)

Obsérvese que la independencia implica que $f_{XY}(x, y) = f_X(x)f_Y(y)$ para *toda* x y y. Si se encuentra un par de x y y para el que la igualdad no sea válida, entonces X y Y no son independientes.

Puede demostrarse que la independencia también implica que $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ para cualesquiera conjuntos A y B.

¿Es rectangular el rango de (X, Y)?

Si el conjunto de puntos en el espacio bidimensional que reciben probabilidades positivas bajo $f_{XY}(x, y)$ no forman un rectángulo, entonces X y Y no son independientes, porque el conocimiento de X puede restringir el rango de valores de Y que reciben una probabilidad positiva. En el ejemplo 6-1 se ilustra este caso. Saber que $X = 3$ implica que Y sólo puede ser igual a 0 o 1. Por consiguiente, la distribución de probabilidad marginal de Y no es igual a la distribución de probabilidad condicional $f_{Y|3}(y)$ para $X = 3$. Utilizando esta idea, se sabe de inmediato que las variables aleatorias X y Y con la función de masa de probabilidad conjunta de la figura 6-1 no son independientes. Si el conjunto de puntos en el espacio bidimensional que recibe probabilidades positivas bajo $f_{XY}(x, y)$ forma un rectángulo, entonces la independencia es posible pero no está demostrada. Aún falta por verificar una de las condiciones.

En vez de verificar la independencia a partir de una distribución de probabilidad conjunta, suele usarse el conocimiento del experimento aleatorio para suponer que dos variables aleatorias son independientes. Después, la función de masa de probabilidad conjunta de X y Y se calcula a partir del producto de las funciones de masa de probabilidad marginal.

EJEMPLO 6-9

En un gran embarque de piezas, 1% de las mismas no cumplen con las especificaciones. El proveedor inspecciona una muestra aleatoria de 30 piezas, y la variable aleatoria X denota el número de piezas de la muestra que no cumplen con las especificaciones. El comprador inspecciona otra muestra aleatoria de 20 piezas, y la variable aleatoria Y denota el número de piezas de esta muestra que no cumple con las especificaciones. ¿Cuál es la probabilidad de que $X \leq 1$ y $Y \leq 1$?

Aun cuando las muestras se seleccionan sin reemplazo, si el embarque es grande, en comparación con el tamaño de las muestras seleccionadas, es posible calcular probabilidades aproximadas suponiendo que el muestreo se hace con reemplazo y que X y Y son independientes. Con este supuesto, la distribución de probabilidad marginal de X es binomial con $n = 30$ y $p = 0.01$, y la distribución de probabilidad marginal de Y es binomial con $n = 20$ y $p = 0.01$.

La probabilidad pedida es

$$\begin{aligned} P(X \leq 1, Y \leq 1) &= P(X = 0, Y = 0 + P(X = 1, Y = 0) \\ &\quad + P(X = 0, Y = 1) + P(X = 1, Y = 1) \\ &= f_{XY}(0, 0) + f_{XY}(1, 0) + f_{XY}(0, 1) + f_{XY}(1, 1) \end{aligned}$$

Por la independencia de X y Y

$$\begin{aligned} P(X \leq 1, Y \leq 1) &= f_X(0) f_Y(0) + f_X(1) f_Y(0) + f_X(0) f_Y(1) + f_X(1) f_Y(1) \\ &= [(0.99)^{30}][(0.99)^{20}] \\ &\quad + [30(0.01)^1(0.99)^{29}][(0.99)^{20}] \\ &\quad + [(0.99)^{30}][20(0.01)^1(0.99)^{19}] \\ &\quad + [30(0.01)^1(0.99)^{29}][20(0.01)^1(0.99)^{19}] \\ &= 0.948 \end{aligned}$$

Además, las distribuciones binomiales pueden usarse para demostrar que $P(X \leq 1, Y \leq 1) = P(X \leq 1)P(Y \leq 1)$, ya que X y Y son independientes.

Si en cada inspección se deja margen para que una pieza no cumpla con las especificaciones, la probabilidad de que el embarque sea aceptado para usarlo en el proceso de manufactura es 0.948 incluso si el 1% de las piezas no cumplen con las especificaciones. Si el proveedor y el comprador cambian sus políticas de tal modo que el embarque sólo se acepte si no se encuentra ninguna pieza en la muestra que no cumpla con las especificaciones, entonces la probabilidad de que el embarque sea aceptado para la producción sigue siendo muy alta. Es decir,

$$P(X = 0, Y = 0) = P(X = 0, Y = 0) = 0.605$$

Con este ejemplo se demuestra que la inspección no es un método eficaz para conseguir la calidad.

EJERCICIOS DE LA SECCIÓN 6-1

6-1. Demuestre que la siguiente función satisface las propiedades de una función de masa de probabilidad conjunta:

x	y	$f_{XY}(x, y)$
1.5	2	1/8
1.5	3	1/4
2.5	4	1/2
3	5	1/8

6-2. Determine las siguientes probabilidades del ejercicio 6-1:
a) $P(X < 2.5, Y < 3)$
b) $P(X < 2.5)$
c) $P(Y < 3)$
d) $P(X > 1.8, Y > 4.7)$

6-3. Determine $E(X)$ y $E(Y)$ del ejercicio 6-1.

6-4. En el ejercicio 6-1, determine
a) la distribución de probabilidad marginal de la variable aleatoria X.
b) la distribución de probabilidad condicional de Y, dado que $X = 1.5$.
c) la distribución de probabilidad condicional de X, dado que $Y = 2$.
d) $E(Y \mid X = 1.5)$

6-5. Determine el valor de c que hace que la función $f(x, y) = c(x + y)$ sea una función de masa de probabilidad conjunta en los nueve puntos con $x = 1, 2, 3$ y $y = 1, 2, 3$.

6-6. Determine las siguientes probabilidades del ejercicio 6-5:
a) $P(X = 1, Y < 4)$
b) $P(X = 1)$
c) $P(Y = 2)$
d) $P(X < 2, Y < 2)$

6-7. Determine $E(X)$, $E(Y)$, $V(X)$ y $V(Y)$ del ejercicio 6-5.

6-8. En el ejercicio 6-5, determine
a) la distribución de probabilidad marginal de la variable aleatoria X.
b) la distribución de probabilidad condicional de Y, dado que $X = 1$.
c) la distribución de probabilidad condicional de X, dado que $Y = 2$.
d) $E(Y \mid X = 1)$

6-9. Demuestre que la siguiente función satisface las propiedades de una función de masa de probabilidad conjunta:

x	y	$f_{XY}(x, y)$
–1	–2	1/8
–0.5	–1	1/4
0.5	1	1/2
1	2	1/8

6-10. Determine las siguientes probabilidades del ejercicio 6-9:
a) $P(X < 0.5, Y < 1.5)$
b) $P(X < 0.5)$
c) $P(Y < 1.5)$
d) $P(X > 0.25, Y < 4.5)$

6-11. Determine $E(X)$ y $E(Y)$ del ejercicio 6-9.

6-12. En el ejercicio 6-9, determine

a) la distribución de probabilidad marginal de la variable aleatoria X.

b) la distribución de probabilidad condicional de Y, dado que $X = 1$.

c) la distribución de probabilidad condicional de X, dado que $Y = 1$.

d) $E(X \mid y = 1)$

6-13. Se seleccionan cuatro impresoras electrónicas de un lote grande de impresoras dañadas. Cada impresora se inspecciona y se clasifica de acuerdo a si la falla que presenta es importante o secundaria. Sea que las variables aleatorias X y Y denoten el número de impresoras con fallas importantes y secundarias, respectivamente, determine el rango de la distribución de probabilidad conjunta de X y Y.

6-14. En la transmisión de información digital, la probabilidad de que un bit presente una distorsión alta, moderada y baja es 0.01, 0.04 y 0.95, respectivamente. Suponga que se transmiten tres bits y que la distorsión mostrada por cada bit es independiente. Sea que X y Y denoten el número de bits, de los tres transmitidos, con distorsión alta y moderada, respectivamente, determine

a) $f_{XY}(x, y)$

b) $f_X(x)$

c) $f_{Y|1}(y)$

d) $E(Y|X = 1)$

6-15. En la fabricación de cinta magnética, un rollo de cinta de 24 pulgadas se corta y embobina en 48 carretes de media pulgada. Sea que las variables aleatorias X y Y denoten el número de carretes defectuosos en un rollo fabricado por un proveedor local y uno internacional, respectivamente. Suponga que el número de carretes defectuosos de los dos proveedores es independiente y que la proporción de carretes defectuosos del proveedor local y del internacional son 2% y 3%, respectivamente.

a) Describa el rango de la distribución de probabilidad conjunta de X y Y.

b) ¿Cuál es la distribución de probabilidad marginal de X? ¿Y la de Y?

c) ¿Cuál es $P(X = 0, Y = 0)$?

d) ¿Cuál es la probabilidad de que tanto X como Y sean menores o iguales que uno?

6-16. Continuación del ejercicio 6-15. ¿Cuáles son $E(X)$, $V(X)$, $E(Y)$ y $V(Y)$?

6-17. Continuación del ejercicio 6-1. ¿Las variables aleatorias del ejercicio 6-1 son independientes? ¿Por qué sí o por qué no?

6-18. Continuación del ejercicio 6-5. ¿Las variables aleatorias del ejercicio 6-5 son independientes? ¿Por qué sí o por qué no?

6-19. Continuación del ejercicio 6-9. ¿Las variables aleatorias del ejercicio 6-9 son independientes? ¿Por qué sí o por qué no?

6-20. Continuación del ejercicio 6-14. ¿Las variables aleatorias del ejercicio 6-14 son independientes? ¿Por qué sí o por qué no?

6-2 VARIABLES ALEATORIAS DISCRETAS MÚLTIPLES

6-2.1 Distribuciones de probabilidad conjunta

En ocasiones, en un experimento aleatorio se definen más de dos variables aleatorias. Por ejemplo, supóngase que la calidad de cada bit recibido en el ejemplo 6-1 se clasifica con mayor detalle en una de las cuatro clases siguientes: excelente, bueno, regular y pobre, denotadas por E, G, F y P, respectivamente.

Sea que las variables aleatorias X_1, X_2, X_3 y X_4 denoten el número de bits que son E, G, F y P, respectivamente, en una transmisión de 20 bits. En este ejemplo nos interesa la distribución de probabilidad conjunta de las cuatro variables aleatorias. La distribución de probabilidad conjunta de X_1, X_2, X_3 y X_4 puede especificarse mediante una **función de masa de probabilidad**

conjunta $f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4) = P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4)$. Debido a que cada uno de los 20 bits se clasifica en una de las cuatro clases, únicamente los valores de x_1, x_2, x_3 y x_4 tales que $x_1 + x_2 + x_3 + x_4 = 20$ reciben una probabilidad positiva en la función de masa de probabilidad de X_1, X_2, X_3 y X_4. La función de masa de probabilidad es cero en cualquier otro punto.

Dadas las variables aleatorias discretas $X_1, X_2, X_3, ..., X_p$, la **distribución de probabilidad conjunta** de $X_1, X_2, X_3, ..., X_p$ es una descripción del conjunto de puntos $(x_1, x_2, x_3, ..., x_p)$ en el rango de $X_1, X_2, X_3, ..., X_p$, junto con la probabilidad de cada punto. Una función de masa de probabilidad conjunta es una extensión simple de una función de masa de probabilidad de dos variables.

Definición

La **función de masa de probabilidad conjunta** de $X_1, X_2, X_3, ..., X_p$ es

$$f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) = P(X_1 = x_1, X_2 = x_2, ..., X_p = x_p) \tag{6-8}$$

para todos los puntos $(x_1, x_2, x_3, ..., x_p)$ en el rango de $X_1, X_2, X_3, ..., X_p$.

Una distribución de probabilidad marginal es una extensión simple del resultado para dos variables aleatorias.

Definición

Si $X_1, X_2, X_3, ..., X_p$ son variables aleatorias discretas con función de masa de probabilidad conjunta $f_{X_1X_2...X_p}(x_1, x_2, ..., x_p)$, entonces la **función de masa de probabilidad marginal** de cualquier X_i es

$$\begin{aligned} f_{X_i}(x_i) &= P(X_i = x_i) \\ &= \sum_{R_{x_i}} f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) \end{aligned} \tag{6-9}$$

donde R_{X_i} denota el conjunto de puntos en el rango de $(X_1, X_2, X_3, ..., X_p)$ para los que $X_i = x_i$.

EJEMPLO 6-10

En la figura 6-5 se muestra la distribución de probabilidad conjunta de las tres variables aleatorias X_1, X_2, X_3. El rango se compone de los enteros no negativos con $x_1 + x_2 + x_3 = 3$. La distribución de probabilidad marginal de X_2 se encuentra de la siguiente manera:

$$P(X_2 = 0) = f_{X_1X_2X_3}(3, 0, 0) + f_{X_1X_2X_3}(0, 0, 3) + f_{X_1X_2X_3}(1, 0, 2) + f_{X_1X_2X_3}(2, 0, 1)$$

$$P(X_2 = 1) = f_{X_1X_2X_3}(2, 1, 0) + f_{X_1X_2X_3}(0, 1, 2) + f_{X_1X_2X_3}(1, 1, 1)$$

$$P(X_2 = 2) = f_{X_1X_2X_3}(1, 2, 0) + f_{X_1X_2X_3}(0, 2, 1)$$

$$P(X_2 = 3) = f_{X_1X_2X_3}(0, 3, 0)$$

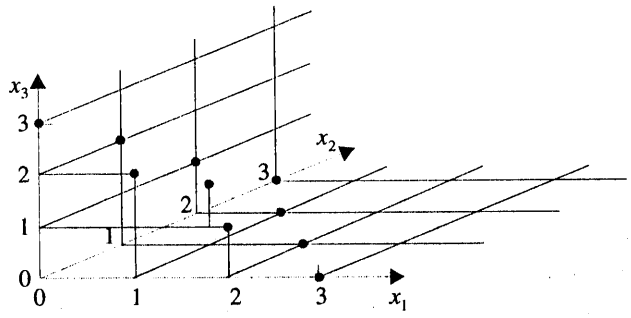


Figura 6-5 Distribución de probabilidad conjunta de X_1, X_2 y X_3.

Además, $E(X_i)$ y $V(X_i)$ para $i = 1, 2, ..., p$ puede determinarse a partir de la distribución de probabilidad marginal de X_i o a partir de la distribución de probabilidad conjunta de $X_1, X_2, ..., X_p$, como se indica a continuación:

$$E(X_i)\sum_R x_i f_{X_1X_2 \ldots X_p}(x_1, x_2 ..., x_p)$$

y

$$V(X_i)=\sum_R (x_i-\mu_{X_i})^2 f_{X_1X_2 \ldots X_p}(x_1, x_2 ..., x_p) \qquad (6\text{-}10)$$

donde R es el conjunto de todos los puntos en el rango de $X_1, X_2, ..., X_p$.

Cuando se trabaja con más de dos variables aleatorias, el interés podría estar en la distribución de probabilidad de un subconjunto de la colección de variables. La distribución de probabilidad de $X_1, X_2, ..., X_k$, con $k < p$, puede obtenerse a partir de la distribución de probabilidad conjunta de $X_1, X_2, ..., X_p$ de la siguiente manera:

Si $X_1, X_2, X_3, ..., X_p$ son variables aleatorias discretas con función de masa de probabilidad conjunta $f_{X_1X_2 \ldots X_p}(x_1, x_2, ... x_p)$, entonces la **función de masa de probabilidad** de $X_1, X_2, ..., X_k$, con $k < p$, es

$$f_{X_1X_2...X_k}(x_1, x_2, ..., x_k)=P(X_1=x_1, X_2=x_2, ..., X_k=x_k)$$
$$= \sum_{R_{x_1x_2...,x_k}} P(X_1=x_1, X_2=x_2, ..., X_k=x_k) \qquad (6\text{-}11)$$

donde $R_{x_1x_2...,x_k}$ denota el conjunto de todos los puntos posibles en el rango de X_1, $X_2, ..., X_p$ para los que $X_1 = x_1, X_2 = x_2, ..., X_k = x_k$.

Es decir, $P(X_1 = x_1, X_2 = x_2, ..., X_k = x_k)$ es la suma de las probabilidades en todos los puntos del rango de $X_1, X_2, X_3, ..., X_p$ para los que $X_1 = x_1, X_2 = x_2, ...,$ y $X_k = x_k$. Más adelante se presenta un ejemplo.

Distribuciones de probabilidad condicional

Es posible desarrollar distribuciones de probabilidad condicional para variables aleatorias discretas múltiples mediante una extensión de las ideas usadas para dos variables aleatorias discretas. Por ejemplo, la función de masa de probabilidad conjunta condicional de X_1, X_2, X_3, dadas X_4, X_5 es

$$f_{X_1X_2X_3|x_4x_5}(x_1, x_2, x_3) = f_{X_1X_2X_3X_4X_5}(x_1, x_2, x_3, x_4, x_5) / f_{X_4X_5}(x_4, x_5)$$

para $f_{X_4X_5}(x_4, x_5) > 0$.

La función de masa de probabilidad conjunta condicional de X_1, X_2, X_3, dadas X_4, X_5 proporciona las probabilidades condicionales en todos los puntos del rango X_1, X_2, X_3, X_4, X_5 para los que $X_4 = x_4$ y $X_5 = x_5$.

El concepto de independencia puede ampliarse a variables aleatorias discretas múltiples.

Definición

Las variables discretas $X_1, X_2, ..., X_p$ son **independientes** si y sólo si

$$f_{X_1X_2...X_p}(x_1, x_2 ..., x_p) = f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_p}(x_p) \qquad (6\text{-}12)$$

para *toda* $x_1, x_2, ..., x_p$.

Como en el resultado obtenido para variables aleatorias de dos variables, la independencia implica que la ecuación 6-12 es válida para *toda* $x_1, x_2, ..., x_p$. Si se encuentra un punto para el que la igualdad no es válida, entonces $X_1, X_2, ..., X_p$ no son independientes. Puede demostrarse que si $X_1, X_2, ..., X_p$ son independientes, entonces $P(X_1 \in A_1, X_2 \in A_2, ..., X_p \in A_p) = P(X_1 \in A_1)P(X_2 \in A_2), ..., P(X_p \in A_p)$ para cualesquiera conjuntos $A_1, A_2, ..., A_p$.

6-2.2 Distribución de probabilidad multinomial

Una distribución de probabilidad conjunta para variables aleatorias discretas múltiples que resulta de suma utilidad, es una extensión de la binomial. El experimento aleatorio que genera la distribución de probabilidad consiste en una serie de ensayos independientes. Sin embargo, los resultados de cada ensayo pueden clasificarse en una de k clases.

EJEMPLO 6-11

Podría haber interés en una probabilidad como la siguiente: de los 20 bits recibidos, ¿cuál es la probabilidad de que 14 sean excelentes, 3 sean buenos, 2 sean regulares y 1 sea pobre? Suponga que la clasificación de los bits individuales son eventos independientes y que las probabilidades de E, G, F y P son 0.6, 0.3, 0.08 y 0.02, respectivamente. Una secuencia de 20 bits que produce el número especificado de bits en cada clase puede representarse como

$$EEEEEEEEEEEEEEGGGFFP$$

Utilizando la independencia, se encuentra que la probabilidad de esta secuencia es

$$P(EEEEEEEEEEEEEEGGGFFP) = 0.6^{14}0.3^{3}0.08^{2}0.02^{1} = 2.708 \times 10^{-9}$$

Evidentemente, todas las secuencias que constan del mismo número de elementos E, G, F y P tienen la misma probabilidad. Por consiguiente, la probabilidad pedida puede encontrarse multiplicando 2.708×10^{-9} por el número de secuencias con 14 E, tres G, dos F y una P. En el apéndice I se encuentra que el número de secuencias es

$$\frac{20!}{14!\,3!\,2!\,1!} = 2\,325\,600$$

Por lo tanto, la probabilidad pedida es

$$\begin{aligned} P(14\ E, \text{tres } G, \text{dos } F, \text{y una } P) &= 2\,325\,600(2.708 \times 10^{-9}) \\ &= 0.0063 \end{aligned}$$

El ejemplo 6-11 lleva a la siguiente generalización de un experimento binomial y una distribución binomial.

Distribución multinomial

Suponga que un experimento aleatorio consta de una serie de n ensayos. Suponga que

1) el resultado de cada ensayo se clasifica en una de k clases;
2) la probabilidad de que un ensayo genere un resultado en la clase 1, en la clase 2, ..., en la clase k es constante en todos los ensayos y es igual a $p_1, p_2, \ldots, p_k$, respectivamente;
3) los ensayos son independientes.

Las variables aleatorias $X_1, X_2, \ldots, X_k$ que denotan el número de ensayos que resultan en la clase 1, en la clase 2, ..., en la clase k, respectivamente, tienen una **distribución multinomial** y la función de masa de probabilidad conjunta es

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1!\,x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k} \tag{6-13}$$

para $x_1 + x_2 + \ldots + x_k = n$ y $p_1 + p_2 + \ldots + p_k = 1$.

La distribución multinomial se considera una extensión para variables múltiples de la distribución binomial.

EJEMPLO 6-12

En el ejemplo 6-11, sea que las variables aleatorias X_1, X_2, X_3 y X_4 denoten el número de bits que son E, G, F y P, respectivamente, en una transmisión de 20 bits. La probabilidad de que 12 de los bits recibidos sean E, 6 sean G, 2 sean F y 0 sean P es

$$\begin{aligned} &P(X_1 = 12, X_2 = 6, X_3 = 2, X_4 = 0) \\ &= \frac{20!}{12!\,6!\,2!\,0!} 0.6^{12} 0.3^{6} 0.08^{2} 0.02^{0} \\ &= 0.0358 \end{aligned}$$

Cada ensayo de un experimento aleatorio multinomial puede considerarse que genera o no un resultado en la clase i, para cada $i = 1, 2, ..., k$. Puesto que la variable aleatoria X_i es el número de ensayos que resultan en la clase i, X_i tiene una distribución binomial.

Si $X_1, X_2, ..., X_k$ tienen una distribución multinomial, la distribución de probabilidad marginal de X_i es binomial con

$$E(X_i) = np_i \quad \text{y} \quad V(X_i) = np_i(1 - p_i) \tag{6-14}$$

EJEMPLO 6-13

En el ejemplo 6-12, la distribución de probabilidad marginal de X_2 es binomial con $n = 20$ y $p = 0.3$. Además, la distribución de probabilidad conjunta marginal de X_2 y X_3 se encuentra como se indica a continuación. La probabilidad $P(X_2 = x_2, X_3 = x_3)$ es la probabilidad de que exactamente x_2 ensayos den como resultado G y de que x_3 dé como resultado F. Los $n - x_2 - x_3$ ensayos restantes tienen que dar como resultado E o bien P. Por consiguiente, puede considerarse que cada ensayo del experimento da como resultado una de tres clases, $\{G\}$, $\{F\}$ o $\{E, P\}$, con probabilidades 0.3, 0.08 y 0.6 + 0.02 = 0.62, respectivamente. Con estas nuevas clases, puede considerarse que los ensayos componen un nuevo experimento multinomial. Por lo tanto,

$$\begin{aligned} f_{X_2X_3}(x_2, x_3) &= P(X_2 = x_2, X_3 = x_3) \\ &= \frac{n!}{x_2!\,x_3!\,(n - x_2 - x_3)!} (0.3)^{x_2} (0.08)^{x_3} (0.62)^{n - x_2 - x_3} \end{aligned}$$

La distribución de probabilidad conjunta de otros conjuntos de variables puede encontrarse de manera similar.

EJERCICIOS DE LA SECCIÓN 6-2

6-21. Suponga que las variables aleatorias X, Y y Z tienen la siguiente distribución de probabilidad conjunta:

x	y	z	$f(x, y, z)$
1	1	1	0.05
1	1	2	0.10
1	2	1	0.15
1	2	2	0.20
2	1	1	0.20
2	1	2	0.15
2	2	1	0.10
2	2	2	0.05

Determine lo siguiente:

a) $P(X = 1)$
b) $P(X = 1, Y = 2)$
c) $P(Z < 1.5)$
d) $P(X = 1 \text{ o } Z = 2)$
e) $E(X)$

6-22. Continuación del ejercicio 6-21. Determine lo siguiente:

a) $P(X = 1 \mid Y = 2)$
b) $P(X = 1, Y = 1 \mid Z = 2)$
c) $P(X = 1 \mid Y = 2, Z = 2)$

6-23. Continuación del ejercicio 6-21. Determine la distribución de probabilidad condicional de X, dado que $Y = 1$ y $Z = 2$.

6-24. Con base en el número de huecos, una placa de ferrita se clasifica como alta, media o baja. Históricamente, 5% de las placas se clasifican como altas, 80% como medias y 15% como bajas. Se selecciona una muestra de prueba de 20 placas. Sea que X, Y y Z denoten el número de placas que se clasifican de manera independiente como altas, medias y bajas, respectivamente.

a) ¿Cuáles son los nombres y los valores de los parámetros de la distribución de probabilidad conjunta de X, Y y Z?
b) ¿Cuál es el rango de la distribución de probabilidad conjunta de X, Y y Z?
c) ¿Cuáles son los nombres y los valores de los parámetros de la distribución de probabilidad marginal de X?
d) Determine $E(X)$ y $V(X)$.

6-25. Continuación del ejercicio 6-24. Determine lo siguiente:

a) $P(X = 1, Y = 17, Z = 3)$
b) $P(X \leq 1, Y = 17, Z = 3)$
c) $P(X \leq 1)$

6-26. Continuación del ejercicio 6-24. Determine lo siguiente:

a) $P(X = 2, Z = 3 \mid Y = 17)$
b) $P(X = 2 \mid Y = 17)$
c) $E(X \mid Y = 17)$

6-27. Un pedido de 15 impresoras contiene cuatro con la característica de gráficas mejoradas, cinco con memoria extra y seis con ambas características. Se seleccionan cuatro impresoras al azar y sin reemplazo de este conjunto. Sea que las variables aleatorias X, Y y Z denoten el número de impresoras con gráficas mejoradas, memoria extra y con ambas características, respectivamente.

a) Describa el rango de la distribución de probabilidad conjunta de X, Y y Z.
b) ¿La distribución de probabilidad de X, Y y Z es multinomial? ¿Por qué sí o por qué no?
c) Determine la distribución de probabilidad condicional de X, dado que $Y = 2$.

6-28. Continuación del ejercicio 6-27. Determine lo siguiente:

a) $P(X = 1, Y = 2, Z = 1)$
b) $P(X = 1, Y = 1)$
c) $E(X)$ y $V(X)$.

6-29. Continuación del ejercicio 6-27. Determine lo siguiente:

a) $P(X = 1, Y = 2 \mid Z = 1)$
b) $P(X = 2 \mid Y = 2)$
c) Determine la distribución de probabilidad condicional de X, dado que $Y = 0$ y $Z = 3$.

6-30. Una muestra de cuatro hornos electrónicos que se dejaron caer durante el transporte se ins-inspeccionan y clasifican de acuerdo a si presentan defectos mayores, menores o sin defectos. En el pasado, 60% de los hornos que se dejaron caer tuvieron un defecto mayor, 30% tuvieron defectos menores y 10% no tuvieron

defectos. Suponga que los defectos en los cuatro hornos ocurren de manera independiente.

a) ¿Cuál es la probabilidad de que, de los cuatro hornos que se dejaron caer, dos tengan un defecto mayor y dos tengan un defecto menor?

b) ¿Cuál es la probabilidad de que ninguno de los hornos tenga un defecto?

6-31. Continuación del ejercicio 6-30. Determine la función de masa de probabilidad conjunta del número de hornos con un defecto mayor y el número con un defecto menor.

6-32. En la transmisión de información digital, la probabilidad de que un bit presente una distorsión alta, moderada o baja es 0.01, 0.04 y 0.95, respectivamente. Considere la transmisión de tres bits en la que se supone independiente la cantidad de distorsión presente en cada bit.

a) ¿Cuál es la probabilidad de que dos bits presenten distorsión alta y uno presente distorsión moderada?

b) ¿Cuál es la probabilidad de que los tres bits presenten distorsión baja?

6-33. Continuación del ejercicio 6-32. Sea que X denote el número de bits con distorsión alta de los tres transmitidos. ¿Cuál es la distribución de probabilidad, la media y la varianza de X?

6-3 DOS VARIABLES ALEATORIAS CONTINUAS

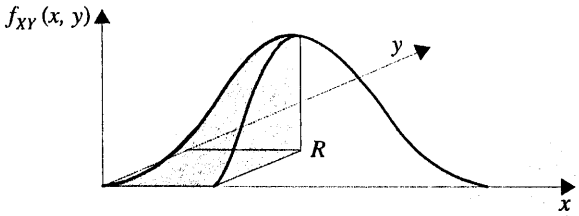
6-3.1 Distribuciones de probabilidad conjunta

La exposición de la distribución de probabilidad conjunta de dos variables aleatorias continuas es similar a la exposición de dos variables aleatorias discretas. Como ejemplo, sea que la variable aleatoria continua X denote la longitud de una dimensión de una pieza moldeada por inyección, y sea que la variable aleatoria continua Y denote la longitud de otra dimensión. El espacio muestral del experimento aleatorio consta de los puntos en dos dimensiones.

Cada variable aleatoria puede estudiarse por separado. Sin embargo, puesto que ambas variables aleatorias son mediciones de la misma pieza, pequeños imprevistos en el proceso de moldeo por inyección, como variaciones en la presión y la temperatura, tendrían mayores posibilidades de generar valores de X y Y en regiones específicas del espacio bidimensional. Por ejemplo, un pequeño aumento de la presión podría generar piezas tales que, tanto X como Y, sean mayores que sus respectivos valores especificados, y pequeños decrementos en la presión podrían generar piezas tales que X y Y sean ambas menores que sus respectivos valores especificados. Por lo tanto, con base en las variaciones de la presión, cabría esperar que sea reducida la probabilidad de una pieza con X mucho mayor que su valor especificado y Y mucho menor que su valor especificado. Conocer la distribución de probabilidad conjunta de X y Y proporciona información que no es evidente en las distribuciones de probabilidad marginal.

La distribución de probabilidad conjunta de dos variables aleatorias continuas X y Y puede especificarse estableciendo un método para calcular la probabilidad de que X y Y asuman un valor en cualquier región R del espacio bidimensional. En forma análoga a la función de densidad de probabilidad de una sola variable aleatoria continua, puede definirse una **función de densidad de probabilidad conjunta** en el espacio bidimensional. La integral doble de $f_{XY}(x, y)$ en una región R proporciona la probabilidad de que (X, Y) asuma un valor en R. Esta integral puede interpretarse como el volumen bajo la superficie $f_{XY}(x, y)$ en la región R.

En la figura 6-6 se muestra una función de densidad de probabilidad conjunta de X y Y. La probabilidad de que (X, Y) asuma un valor en la región R es igual al volumen de la región sombreada de la figura 6-6. De este modo, se utiliza una función de densidad de probabilidad conjunta para determinar las probabilidades de X y Y.



La probabilidad de que (X, Y) esté en la región R se determina por el volumen de $f_{XY}(x, y)$ en la región R.

Figura 6-6 Función de densidad de probabilidad conjunta de las variables aleatorias X y Y.

Definición

Una **función de densidad de probabilidad conjunta** de las variables aleatorias continuas X y Y, denotada como $f_{XY}(x, y)$, satisface las siguientes propiedades:

1) $f_{XY}(x, y) \geq 0$ para toda x, y

2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\,dx\,dy = 1$

3) para cualquier región R del espacio bidimensional,

$$P([X, Y] \in R) = \iint_R f_{XY}(x, y)\,dx\,dy \qquad (6\text{-}15)$$

De manera típica, $f_{XY}(x, y)$ se define en todo el espacio bidimensional suponiendo que $f_{XY}(x, y) = 0$ para todos los puntos en los que $f_{XY}(x, y)$ no está especificada.

Al principio de este capítulo se presentaron como ejemplo de dos variables aleatorias las longitudes de diferentes dimensiones de una pieza moldeada por inyección. Cada longitud podía modelarse con una distribución normal. Sin embargo, puesto que las mediciones se hacen en la misma pieza, las variables aleatorias por lo general no son independientes. Una distribución de probabilidad de dos variables aleatorias normales que no son independientes es importante en muchas aplicaciones y se presenta más adelante en este capítulo. Si las especificaciones de X y Y son (2.95, a 3.05) y (7.60 a 7.80) milímetros, respectivamente, entonces podríamos interesarnos en la probabilidad de que una pieza satisfaga ambas especificaciones; es decir, $P(2.95 < X < 3.05, 7.60 < Y < 7.80)$. Suponga que $f_{XY}(x, y)$ se muestra en la figura 6-7. La probabilidad buscada es el volumen de $f_{XY}(x, y)$ dentro de las especificaciones. Con frecuencia, una probabilidad como ésta debe determinarse por integración numérica.

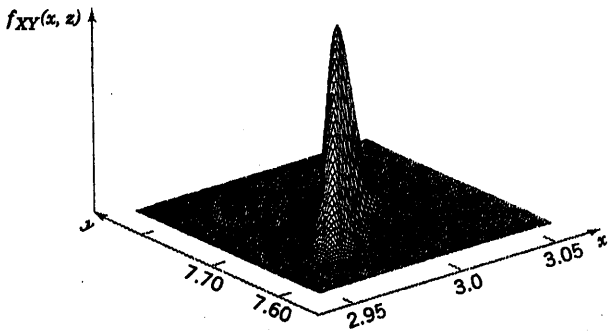


Figura 6-7 Función de densidad de probabilidad conjunta para las longitudes de diferentes dimensiones de una pieza moldeada por inyección.

EJEMPLO 6-14

Para una copiadora, sea que la variable aleatoria X denote el tiempo hasta la falla de un componente en horas, y sea que Y represente el tiempo hasta la falla de un componente de repuesto en horas. Cada una de estas variables aleatorias mide el tiempo hasta una falla a partir de un tiempo inicial común. Puesto que el repuesto no se pone en uso hasta que falla el primer componente, $X < Y$. Suponga que la vida útil del componente y del repuesto tiene una distribución exponencial con parámetros 0.003 y 0.002, respectivamente. (La función de densidad de probabilidad conjunta de X y Y puede deducirse por el método de transformación que se expone en el apéndice III. Sin embargo, esto se sale del alcance de la presente exposición, y el resultado se establece directamente.) La función de densidad de probabilidad conjunta de X y Y es

$$f_{XY}(x, y) = 6 \times 10^{-6} \quad \exp \quad (-0.001x - 0.002y) \quad \text{para} \quad x < y$$

Puesto que el repuesto no se pone en uso hasta que falla el primer componente, su tiempo hasta una falla no es una distribución exponencial. La distribución de Y se establece más adelante en el capítulo. Por el momento, la atención se centra en la función de densidad de probabilidad conjunta.

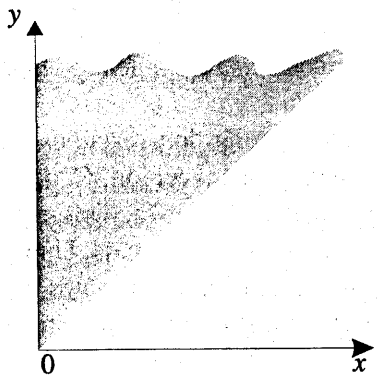


Figura 6-8 La función de densidad de probabilidad conjunta de X y Y del ejemplo 6-14 es diferente de cero en la región sombreada.

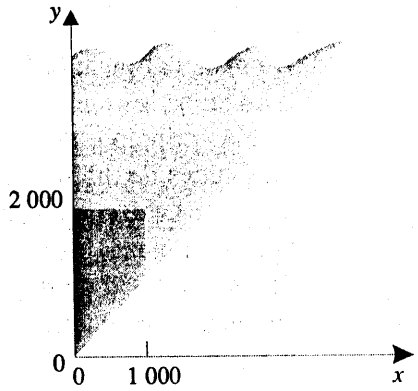


Figura 6-9 La región de integración para la probabilidad de que $X < 1\,000$ y $Y < 2\,000$ se indica con el sombreado más oscuro.

La región con probabilidad diferente de cero está sombreada en la figura 6-8. La propiedad de que la integral de esta función de densidad de probabilidad conjunta es uno puede comprobarse con la integral de $f_{XY}(x, y)$ en esta región de la siguiente manera:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\,dy\,dx = \int_{0}^{\infty}\left(\int_{x}^{\infty} 6\times 10^{-6} e^{-0.001x-0.002y}dy\right)dx$$

$$= 6\times 10^{-6}\int_{0}^{\infty}\left(\int_{x}^{\infty} e^{-0.002y}dy\right)e^{-0.001x}dx$$

$$= 6\times 10^{-6}\int_{0}^{\infty}\left(\frac{e^{-0.002x}}{0.002}\right)e^{-0.001x}dx$$

$$= \frac{6\times 10^{-6}}{0.002}\left(\int_{0}^{\infty} e^{-0.003x}dx\right) = \frac{6\times 10^{-6}}{0.002}\left(\frac{1}{0.003}\right) = 1$$

La probabilidad de que $X < 1\,000$ y $Y < 2\,000$ se determina como la integral en la región con el sombreado más oscuro de la figura 6-9.

$$P(X \le 1\,000, Y \le 2\,000) = \int_{0}^{1\,000}\int_{0}^{2\,000} f_{XY}(x, y)\,dy\,dx$$

$$= 6\times 10^{-6}\int_{0}^{1\,000}\left(\int_{x}^{2\,000} e^{-0.002y}dy\right)e^{-0.001x}dx$$

$$= 6\times 10^{-6}\int_{0}^{1\,000}\left(\frac{e^{-0.002x}-e^{-4}}{0.002}\right)e^{-0.001x}dx$$

$$= \frac{6\times 10^{-6}}{0.002}\int_{0}^{1\,000} e^{-0.003x} - e^{-4}e^{-0.001x}dx$$

$$= 3\times 10^{-3}\left[\left(\frac{1-e^{-3}}{0.03}\right) - e^{-4}\left(\frac{1-e^{-1}}{0.001}\right)\right]$$

$$= 3\times 10^{-3}(316.738 - 11.578) = 0.915$$

6-3.2 Distribuciones de probabilidad marginal

En forma similar a las variables aleatorias discretas conjuntas, las distribuciones de probabilidad marginal de X y Y pueden encontrarse a partir de la distribución de probabilidad conjunta.

Definición

Si la función de densidad de probabilidad conjunta de las variables aleatorias continuas X y Y es $f_{XY}(x, y)$, entonces las **funciones de densidad de probabilidad marginal** de X y Y son

$$f_X(x) = \int_{R_x} f_{XY}(x, y)\, dy$$

$$f_Y(y) = \int_{R_y} f_{XY}(x, y)\, dy \qquad (6\text{-}16)$$

donde

R_x denota el conjunto de todos los puntos en el rango de (X, Y) para los que $X = x$ y $4R_y$ denota el conjunto de todos los puntos en el rango de (X, Y) para los que $Y = y$.

Una probabilidad en la que interviene una sola variable aleatoria, por ejemplo $P(a < X < b)$, puede encontrarse a partir de la distribución de probabilidad marginal de X, o bien a partir de la distribución de probabilidad conjunta de X y Y. Por ejemplo, $P(a < X < b)$ es igual a $P(a < X < b, -\infty < Y < \infty)$. Por lo tanto,

$$P(a < X < b) = \int_a^b \int_{R_x} f_{XY}(x, y)\, dy\, dx$$

$$= \int_a^b \left(\int_{R_x} f_{XY}(x, y) dy \right) dx$$

$$= \int_a^b f_X(x)\, dx$$

EJEMPLO 6-15

Para las variables aleatorias que denotan la vida útil del ejemplo 6-14, encuentre la probabilidad de que el tiempo hasta la falla de Y exceda 2 000 horas.

Esta probabilidad se determina como la integral de $f_{XY}(x, y)$ en la región con el sombreado más oscuro de la figura 6-10. La región se divide en dos partes y se determinan límites de integración diferentes para cada parte.

$$P(Y > 2\,000) = \int_0^{2\,000} \left(\int_{2\,000}^{\infty} 6\times10^{-6} e^{-0.001x-0.002y} dy \right) dx$$

$$+ \int_{2\,000}^{\infty} \left(\int_{x}^{\infty} 6\times10^{-6} e^{-0.001x-0.002y} dy \right) dx$$

La primera integral es

$$6\times10^{-6} \int_0^{2\,000} \left(\frac{e^{-0.002y}}{-0.002} \Bigg|_{2000}^{\infty} \right) e^{-0.001x} dx = \frac{6\times10^{-6}}{0.002} e^{-4} \int_0^{2\,000} e^{-0.001x} dx$$

$$= \frac{6\times10^{-6}}{0.002} e^{-4} \left(\frac{1-e^{-2}}{0.001} \right) = 0.0475$$

La segunda integral es

$$6\times10^{-6} \int_{2\,000}^{\infty} \left(\frac{e^{-0.002y}}{-0.002} \Bigg|_{x}^{\infty} \right) e^{-0.001x} dx = \frac{6\times10^{-6}}{0.002} \int_{2\,000}^{\infty} e^{-0.003x} dx$$

$$= \frac{6\times10^{-6}}{0.002} \left(\frac{e^{-6}}{0.003} \right) = 0.0025$$

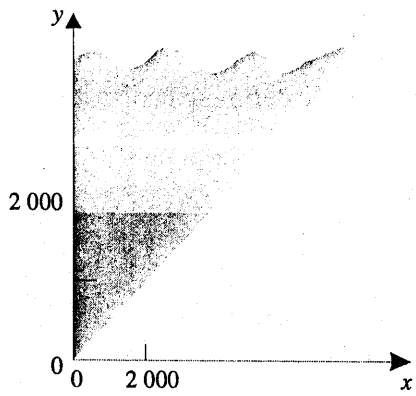


Figura 6-10 La región de integración para la probabilidad de que $Y > 2\,000$ tiene un sombreado más oscuro y se divide en dos regiones con $x < 2\,000$ y $x > 2\,000$.

Por lo tanto,

$$P(Y > 2\,000) = 0.0475 + 0.0025 = 0.05.$$

Alternativamente, la probabilidad puede calcularse a partir de la distribución de probabilidad marginal de Y como se indica a continuación. Para $y > 0$,

$$f_Y(y) = \int_0^y 6\times10^{-6} e^{-0.001x-0.002y} dx = 6\times10^{-6} e^{-0.002y} \int_0^y e^{-0.001x} dx$$

$$= 6\times10^{-6} e^{-0.002y} \left(\frac{e^{-0.001x}}{-0.001} \bigg|_0^y \right) = 6\times10^{-6} e^{-0.002y} \left(\frac{1-e^{-0.001y}}{0.001} \right)$$

$$= 6\times10^{-3} e^{-0.002y} \left(1 - e^{-0.001y}\right) \text{ para } y > 0$$

Se ha obtenido la función de densidad de probabilidad marginal de Y. Ahora,

$$P(Y > 2\,000) = 6\times10^{-3} \int_{2\,000}^{\infty} e^{-0.002y} \left(1 - e^{-0.001y}\right) dy$$

$$= 6\times10^{-3} \left[\left(\frac{e^{-0.002y}}{-0.002} \bigg|_{2\,000}^{\infty} \right) - \left(\frac{e^{-0.003y}}{-0.003} \bigg|_{2\,000}^{\infty} \right) \right]$$

$$= 6\times10^{-3} \left[\frac{e^{-4}}{0.002} - \frac{e^{-6}}{0.003} \right] = 0.05$$

6-3.3 Distribuciones de probabilidad condicional

En forma análoga a las variables aleatorias discretas, puede definirse la distribución de probabilidad condicional de Y, dado que $X = x$.

Definición

Dadas las variables aleatorias continuas X y Y con función de densidad de probabilidad conjunta $f_{XY}(x, y)$, la **función de densidad de probabilidad condicional** de Y, dado que $X = x$, es

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \text{para } f_X(x) > 0 \qquad (6\text{-}17)$$

La función $f_{Y|x}(y)$ se usa para encontrar las probabilidades de los valores posibles de Y, dado que $X = x$. Sea que R_x denote el conjunto de todos los puntos en el rango de (X, Y) para los que $X = x$. La función de densidad de probabilidad condicional proporciona las probabilidades condicionales para los valores de Y en el conjunto R_x.

Puesto que la función de densidad de probabilidad condicional $f_{Y|x}(y)$ es una función de densidad de probabilidad para toda y en R_x, se satisfacen las siguientes propiedades:

1) $f_{Y|x}(y) \geq 0$

2) $\int_{R_x} f_{Y|x}(y)\,dy = 1$

3) $P(Y \in B | X = x) = \int_B f_{Y|x}(y)\,dy$ para cualquier conjunto B en el rango de Y

(6-18)

Es importante enunciar la región en la que una función de densidad de probabilidad conjunta, marginal o condicional, es diferente de cero. En el ejemplo siguiente se ilustra esto.

EJEMPLO 6-16

Para las variables aleatorias que denotan tiempos de vida útil del ejemplo 6-14, determine la función de densidad de probabilidad condicional de Y, dado que $X = x$.

Primero se determina la función de densidad de probabilidad marginal de x. Para $x > 0$

$$f_X(x) = \int_x^\infty 6\times10^{-6} e^{-0.001x-0.002y}\,dy$$

$$= 6\times10^{-6} e^{-0.001x}\left(\frac{e^{-0.002y}}{-0.002}\Bigg|_x^\infty\right)$$

$$= 6\times10^{-6} e^{-0.001x}\left(\frac{e^{-0.002x}}{-0.002}\right) = 0.003e^{-0.003x} \qquad \text{para } x > 0$$

Se trata de una distribución exponencial con $\lambda = 0.003$, como era de esperarse por la descripción de X. Ahora bien, para $0 < x$ y $x < y$, la función de densidad de probabilidad condicional es

$$
\begin{aligned}
f_{Y|X}(y) &= f_{YX}(x, y)/f_x(x) \\
&= \frac{6\times10^{-6}\, e^{-0.001x-0.002y}}{0.003e^{-0.003x}} \\
&= 0.002e^{-0.002x-0.002y} \qquad \text{para } 0 < x \text{ y } x < y
\end{aligned}
$$

La función de densidad de probabilidad condicional de Y, dado que $x = 1\ 500$, es diferente de cero en la recta marcada en la figura 6-11.

Determine la probabilidad de que el tiempo hasta la falla del repuesto exceda 2 000 horas dado que $x = 1\ 500$. Es decir, determine $P(Y > 2\ 000|x = 1\ 500)$. La función de densidad de probabilidad condicional se integra de la siguiente manera:

$$
\begin{aligned}
P(Y > 2\ 000|x = 1\ 500) &= \int_{2000}^{\infty} f_{Y|1\,500}(y)dx \\
&= \int_{2\,000}^{\infty} 0.002e^{0.002(1\,500)-0.002y}\, dy \\
&= 0.002e^{3}\left(\frac{e^{-0.002y}}{-0.002}\Bigg|_{2\,000}^{\infty}\right) = 0.002e^{3}\left(\frac{e^{-4}}{0.002}\right) = 0.368
\end{aligned}
$$

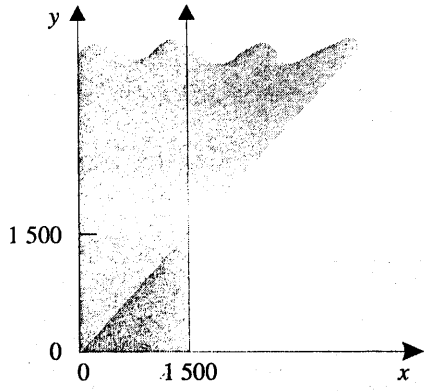


Figura 6-11 La función de densidad de probabilidad condicional para Y, dado que $x = 1\ 500$, en el ejemplo 6-16, es diferente de cero sobre la recta marcada.

Definición

Sea que R_x denote el conjunto de todos los puntos en el rango de (X, Y) para el que $X = x$. La **media condicional** de Y, dado que $X = x$, denotada como $E(Y \mid x)$ o $\mu_{Y|x}$, es

$$E(Y|x) = \int_{R_x} y f_{Y|x}(y)dy$$

y la **varianza condicional** de Y, dado que $X = x$, denotada como $V(Y \mid x)$ o $\sigma^2_{Y|x}$, es

$$V(Y|x) = \int_{R_x} (y - \mu_{Y|x})^2 f_{Y|x}(y)dy$$

$$= \int_{R_x} y^2 f_{Y|x}(y)dy - \mu^2_{Y|x} \qquad (6\text{-}19)$$

EJEMPLO 6-17

Para las variables aleatorias que denotan el tiempo de vida útil en el ejemplo 6-14, determine la media condicional de Y, dado que $x = 1\ 500$.

En el ejemplo 6-16 se determinó la función de densidad de probabilidad condicional de Y. Puesto que $f_{Y|1\ 500}(y)$ es diferente de cero para $y > 1\ 500$

$$E(Y|x = 1\,500) = \int_{1\,500}^{\infty} y(0.002e^{0.002(1\,500)-0.002y})dy$$

$$= 0.002e^3 \int_{1\,500}^{\infty} ye^{-0.002y}dy$$

La integración por partes se hace como sigue:

$$\int_{1\,500}^{\infty} ye^{-0.002y}dy = y\frac{e^{-0.002y}}{-0.002}\Bigg|_{1\,500}^{\infty} - \int_{1\,500}^{\infty}\left(\frac{e^{-0.002y}}{-0.002}\right)dy$$

$$= \frac{1\,500}{0.002}e^{-3} - \left(\frac{e^{-0.002y}}{(-0.002)(-0.002)}\Bigg|_{1\,500}^{\infty}\right)$$

$$= \frac{1\,500}{0.002}e^{-3} + \frac{e^{-3}}{(-0.002)(0.002)} = \frac{e^{-3}}{0.002}(2\,000)$$

Al aplicar de nuevo la constante $0.002e^3$

$$E(Y \mid x = 1\,500) = 2\,000$$

6-3.4 Independencia

La definición de independencia para variables aleatorias continuas es similar a la definición para variables aleatorias discretas. Si $f_{XY}(x, y) = f_X(x)f_Y(y)$ para toda x y y, entonces X y Y son **independientes**.

Para las variables aleatorias continuas X y Y, si cualquiera de las propiedades siguientes es válida, entonces las demás también lo son, y se dice que X y Y son **independientes**.

1) $f_{XY}(x, y) = f_X(x)f_Y(y)$ para toda x y y

2) $f_{Y|x}(y) = f_Y(y)$ para toda x y y con $f_X(x) > 0$

3) $f_{X|y}(x) = f_X(x)$ para toda x y y con $f_Y(y) > 0$

4) $P(X \in A, Y \in B) = P(Y \in A)P(Y \in B)$ para cualesquiera conjuntos A y B en el rango de X y Y, respectivamente. (6-20)

La independencia implica que $f_{XY}(x, y) = f_X(x)f_Y(y)$ para *toda* x y y. Si se encuentra un par de x y y para el que la igualdad no sea válida, entonces X y Y no son independientes. Puede demostrarse que la independencia implica que la propiedad 4 de la ecuación 6-20 es válida para cualesquiera conjuntos A y B (véanse los ejercicios para desarrollar el intelecto).

EJEMPLO 6-18

Para la distribución conjunta del tiempo de falla del ejemplo 6-14,

- la distribución marginal de Y se determinó en el ejemplo 6-15 y
- la distribución condicional de Y, dado que $X = x$, se determinó en el ejemplo 6-16.

Debido a que las densidades de probabilidad condicional y marginal no son las mismas para todos los valores de x, la propiedad 2 de la ecuación 6-20 implica que las variables aleatorias no son independientes. El hecho de que estas variables no son independientes puede determinarse de inmediato al observar que el rango de (X, Y), indicado en la figura 6-8, no es rectangular. Por consiguiente, el conocimiento de X cambia el intervalo de los valores de Y que reciben una probabilidad diferente de cero.

EJEMPLO 6-19

Suponga que el ejemplo 6-14 se modifica de tal modo que la función de densidad de probabilidad conjunta de X y Y es $f_{XY}(x, y) = 2 \times 10^{-6}e^{-0.001x - 0.002y}$ para $x \geq 0$ y $y \geq 0$. Demuestre que X y Y son independientes y determine $P(X > 1\,000, Y < 1\,000)$.

La función de densidad de probabilidad marginal de X es

$$f_X(x) = \int_0^\infty 2\times10^{-6} e^{-0.001x-0.002y} dy$$

$$= 0.001e^{-0.001x} \quad \text{para } x > 0$$

La función de densidad de probabilidad marginal de y es

$$f_Y(y) = \int_0^\infty 2\times10^{-6} e^{-0.001x-0.002y} dx$$

$$= 0.002e^{-0.002y} \quad \text{para } y > 0$$

Por lo tanto, $f_{XY}(x, y) = f_X(x)f_Y(y)$ para toda x y y, y X y Y son independientes.

Para determinar la probabilidad pedida, puede aplicarse la propiedad 4 de la ecuación 6-20 y el hecho de que cada variable aleatoria tiene una distribución exponencial.

$$\begin{aligned} P(X > 1\,000,\ Y < 1\,000) &= P(X > 1\,000)P(Y < 1\,000) \\ &= e^{-1}(1 - e^{-2}) \\ &= 0.318 \end{aligned}$$

Frecuentemente, con base en el conocimiento del sistema bajo estudio, se supone que las variables aleatorias son independientes. Entonces, las probabilidades en las que intervienen ambas variables pueden determinarse a partir de las distribuciones de probabilidad marginal.

EJEMPLO 6-20

Sea que las variables aleatorias X y Y denoten las longitudes de dos dimensiones de una pieza maquinada, respectivamente. Suponga que X y Y son variables aleatorias independientes, y suponga asimismo que la distribución de X es normal con una media de 10.5 milímetros y una varianza de 0.0025 (milímetros)^2, y que la distribución de Y es normal con una media de 3.2 milímetros y una varianza de 0.0036 (milímetros)^2. Determine la probabilidad de que $10.4 < X < 10.6$ y $3.15 < Y < 3.25$.

Como X y Y son independientes,

$$P(10.4 < X < 10.6, 3.15 < Y < 3.25) = P(10.4 < X < 10.6)P(3.15 < Y < 3.25)$$

$$= P\left(\frac{10.4-10.5}{0.05} < Z < \frac{10.6-10.5}{0.05}\right)P\left(\frac{3.15-3.2}{0.06} < Z < \frac{3.25-3.2}{0.06}\right)$$

$$= P(-2 < Z < 2)P(-0.833 < Z < 0.833)$$

$$= 0.566$$

donde Z denota una variable aleatoria normal estándar.

EJERCICIOS DE LA SECCIÓN 6-3

6-34. Determine el valor de c, tal que la función $f(x, y) = cxy$ para $0 < x < 3$ y $0 < y < 3$ satisfaga las propiedades de una función de densidad de probabilidad conjunta.

6-35. Continuación del ejercicio 6-34. Determine lo siguiente:
a) $P(X < 2.5, Y < 3)$
b) $P(X < 2.5)$
c) $P(1 < Y < 2.5)$
d) $P(X > 1.8, 1 < Y < 2.5)$
e) $E(X)$
f) $E(Y)$

6-36. Continuación del ejercicio 6-34. Determine lo siguiente:
a) La distribución de probabilidad marginal de la variable aleatoria X.
b) La distribución de probabilidad condicional de Y, dado que $X = 1.5$.
c) $E(Y \mid x = 1.5)$.
d) La distribución de probabilidad condicional de X, dado que $Y = 2$.

6-37. Determine el valor de c que hace que la función $f(x, y) = c(x + y)$ sea una función de densidad de probabilidad conjunta en el rango $0 < x < 3$ y $x < y < x + 2$.

6-38. Continuación del ejercicio 6-37. Determine lo siguiente:
a) $P(X < 1, Y < 2)$
b) $P(1 < X < 2)$
c) $P(Y > 2)$
d) $P(X < 2, Y < 2)$
e) $E(X)$

6-39. Continuación del ejercicio 6-37. Determine lo siguiente:
a) La distribución de probabilidad marginal de X.
b) La distribución de probabilidad condicional de Y, dado que $X = 1$.
c) $E(Y \mid x = 1)$.
d) La distribución de probabilidad condicional de X, dado que $Y = 2$.

6-40. Determine el valor de c que hace que la función $f(x, y) = cxy$ sea una función de densidad de probabilidad conjunta en el rango $0 < x < 3$ y $0 < y < x$.

6-41. Continuación del ejercicio 6-40. Determine lo siguiente:
a) $P(X < 1, Y < 2)$
b) $P(1 < X < 2)$
c) $P(Y > 2)$
d) $P(X < 2, Y < 2)$
e) $E(X)$
f) $E(Y)$

6-42. Continuación del ejercicio 6-40. Determine lo siguiente:
a) La distribución de probabilidad marginal de X.
b) La distribución de probabilidad condicional de Y, dado que $X = 1$.
c) $E(Y \mid x = 1)$.
d) La distribución de probabilidad condicional de X, dado que $Y = 2$.

6-43. Determine el valor de c que hace que la función $f(x, y) = ce^{-2x-3y}$ sea una función de densidad de probabilidad conjunta en el rango $0 < x$ y $0 < y < x$.

6-44. Continuación del ejercicio 6-43. Determine lo siguiente:
a) $P(X < 1, Y < 2)$
b) $P(1 < X < 2)$
c) $P(Y > 2)$
d) $P(X < 2, Y < 2)$
e) $E(X)$
f) $E(Y)$

6-45. Continuación del ejercicio 6-43. Determine lo siguiente:
a) La distribución de probabilidad marginal de X.
b) La distribución de probabilidad condicional de Y, dado que $X = 1$.
c) $E(Y \mid x = 1)$.
d) La distribución de probabilidad condicional de X, dado que $Y = 2$.

6-46. Determine el valor de c que hace que la función $f(x, y) = ce^{-2x-3y}$ sea una función de densidad de probabilidad conjunta en el rango $0 < x$ y $x < y$.

6-47. Continuación del ejercicio 6-46. Determine lo siguiente:
a) $P(X < 1, Y < 2)$

b) $P(1 < X < 2)$
c) $P(Y > 2)$
d) $P(X < 2, Y < 2)$
e) $E(X)$
f) $E(Y)$

6-48. Continuación del ejercicio 6-46. Determine lo siguiente:
a) La distribución de probabilidad marginal de X.
b) La distribución de probabilidad condicional de Y, dado que $X = 1$.
c) $E(Y \mid x = 1)$.
d) La distribución de probabilidad condicional de X, dado que $Y = 2$.

6-49. Se usan dos métodos para medir la suavidad de la superficie de un producto de papel. Las mediciones se registran como desviaciones de la suavidad nominal de la superficie en unidades codificadas. La distribución de probabilidad conjunta de las dos mediciones es una distribución uniforme en la región $0 < x < 4$, $0 < y$, y $x - 1 < y < x + 1$. Es decir, $f_{XY}(x, y) = c$ para x y y en la región. Determine el valor de c tal que $f_{XY}(x, y)$ sea una función de densidad de probabilidad conjunta.

6-50. Continuación del ejercicio 6-49. Determine lo siguiente:
a) $P(X < 0.5, Y < 0.5)$
b) $P(X < 0.5)$
c) $E(X)$
d) $E(Y)$

6-51. Continuación del ejercicio 6-49. Determine lo siguiente:
a) La distribución de probabilidad marginal de X.
b) La distribución de probabilidad condicional de Y, dado que $X = 1$.
c) $E(Y \mid x = 1)$.

6-52. La función de densidad de probabilidad conjunta de la longitud de dos dimensiones de una pieza es $\frac{1}{1.2\pi} e^{\frac{-1}{0.72}\left[(x-1)^2 - 1.6(x-1)(y-2) + (y-2)^2\right]}$ para $-\infty < x < \infty$ y $-\infty < y < \infty$.
Determine lo siguiente:
a) La distribución de probabilidad marginal de X. [*Sugerencia:* Complete el trinomio cuadrado perfecto en el exponente de la integral sobre y y utilice el hecho de que la integral de la función de densidad de probabilidad normal es uno.]
b) $E(X)$

6-53. Continuación del ejercicio 6-52. Determine la distribución de probabilidad condicional de Y, dado que $X = 1$. ¿Cuál es el nombre de esta distribución de probabilidad?

6-54. El tiempo entre los problemas de terminado superficial en un proceso de galvanización tiene una distribución exponencial con una media de 40 horas. En una planta operan tres líneas de galvanizado cuyo funcionamiento se supone independiente.
a) ¿Cuál es la probabilidad de que no se experimenten problemas de terminado superficial en ninguna de las líneas en 40 horas de operación?
b) ¿Cuál es la probabilidad de que en las tres líneas se experimente un problema de terminado superficial dentro de 20 y 40 horas de operación?
c) ¿Por qué no es necesaria la función de densidad de probabilidad conjunta para responder la pregunta anterior?

6-4 VARIABLES ALEATORIAS CONTINUAS MÚLTIPLES

Como en el caso de variables aleatorias discretas, en ocasiones se definen más de dos variables aleatorias continuas en un experimento aleatorio.

EJEMPLO 6-21

Durante la producción de una pieza maquinada se hacen mediciones rutinarias de varias de sus dimensiones. Sea que las variables aleatorias X_1, X_2, X_3 y X_4 denoten las longitudes de cuatro

dimensiones de una pieza. Entonces, hay por lo menos cuatro variables aleatorias de interés en este estudio.

La distribución de probabilidad conjunta de las variables aleatorias continuas $X_1, X_2, X_3, \ldots, X_p$ puede especificarse estableciendo un método para calcular la probabilidad de que $X_1, X_2, X_3, \ldots, X_p$ asuman un valor en la región R del espacio de p dimensiones. Se utiliza una **función de densidad de probabilidad conjunta** $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)$ para determinar la probabilidad de que $(X_1, X_2, X_3, \ldots, X_p)$ asuman un valor en la región R mediante la integral múltiple de $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)$ en la región R.

Definición

Una **función de densidad de probabilidad conjunta** para las variables aleatorias continuas $X_1, X_2, X_3, \ldots, X_p$, denotada como $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)$, satisface las siguientes propiedades

1) $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p) \geq 0$

2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1dx_2\ldots dx_p = 1$

3) para cualquier región B del espacio de p dimensiones,

$$P[(X_1, X_2, \ldots, X_p) \in B] = \int\int_B \cdots\int f_{X_1,X_2,\ldots,X_p}(x_1, x_2, \ldots, x_p)\,dx_1\,dx_2\ldots dx_p \qquad (6\text{-}21)$$

De manera típica, $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)$ está definida en todo el espacio de p dimensiones suponiendo que $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p) = 0$ para todos los puntos en los que $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)$ no está especificada.

EJEMPLO 6-22

En un dispositivo electrónico, sea que las variables aleatorias X_1, X_2, X_3 y X_4 denoten los tiempos de vida útil de cuatro componentes en horas. Suponga que la función de densidad de probabilidad conjunta de estas variables es

$$f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4) = 9\times10^{-2}e^{-0.001x_1-0.002x_2-0.0015x_3-0.003x_4}$$

$$\text{para } x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0$$

¿Cuál es la probabilidad de que el dispositivo funcione más de 1 000 horas sin fallas?

La probabilidad pedida es $P(X_1 > 1\,000, X_2 > 1\,000, X_3 > 1\,000, X_4 > 1\,000)$, que es igual a la integral múltiple de $f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4)$ en la región $x_1 > 1\,000$, $x_2 > 1\,000$, $x_3 > 1\,000$, $x_4 > 1\,000$. La función de densidad de probabilidad conjunta puede escribirse como un producto de funciones exponenciales, y cada integral es la integral simple de una función exponencial. Por lo tanto,

$$P(X_1 > 1\,000, X_2 > 1\,000, X_3 > 1\,000, X_4 > 1\,000) = e^{-1-2-1.5-3}$$
$$= 0.00055$$

Suponga que la función de densidad de probabilidad conjunta de varias variables aleatorias continuas es una constante, digamos c, en una región R (y que es cero fuera de ella). En este caso especial,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1\,dx_2\ldots dx_p$$
$$= c \times (\text{volumen de la región } R) = 1$$

por la propiedad 2 de la ecuación 6-21. Por lo tanto, $c = 1/\text{volumen}\ (R)$. Además, por la propiedad 3 de la ecuación 6-21,

$$P[(X_1, X_2, \ldots X_p) \in B]$$
$$= \int\int_B \cdots\int f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1\,dx_2\ldots dx_p = c \times \text{volumen}\,(B \cap R)$$
$$= \frac{\text{volumen}\,(B \cap R)}{\text{volumen}\,(R)}$$

Cuando la función de densidad de probabilidad conjunta es constante, la probabilidad de que las variables aleatorias asuman un valor en la región B es simplemente la razón del volumen de la región $B \cap R$ y el volumen de la región R para la que la probabilidad es positiva.

EJEMPLO 6-23

Suponga que la función de densidad de probabilidad conjunta de las variables aleatorias continuas X y Y es constante en la región $x^2 + y^2 \leq 4$. Determine la probabilidad de que $X^2 + Y^2 \leq 1$.

La región que recibe una probabilidad positiva es un círculo de radio 2. Por lo tanto, el área de la región es 4π. El área de la región $x^2 + y^2 \leq 1$ es π. Por consiguiente, la probabilidad pedida es $\pi/(4\pi) = 1/4$.

Definición

Si la función de densidad de probabilidad conjunta de las variables aleatorias continuas $X_1, X_2, ..., X_p$ es $f_{X_1X_2...X_p}(x_1, x_2, ..., x_p)$, entonces la **función de densidad de probabilidad marginal** de X_i es

$$f_{X_i}(x_i) = \int \int_{R_{x_i}} ... \int f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) dx_1 dx_2 ... dx_{i-1}\, dx_{i+1} ... dx_p \tag{6-22}$$

donde R_{x_i} denota el conjunto de todos los puntos en el rango de $X_1, X_2, ..., X_p$ para los que $X_i = x_i$.

Como en el caso de dos variables aleatorias, una probabilidad en la que interviene una sola variable aleatoria, por ejemplo, $P(a < X_i < b)$, puede determinarse a partir de la distribución de probabilidad marginal de X_i o a partir de la distribución de probabilidad conjunta de $X_1, X_2, ..., X_p$. Es decir,

$$P(a < X_i < b) = P(-\infty < X_1 < \infty, ... -\infty < X_{i-1} < \infty, a < X_i < b$$
$$-\infty < X_{i+1} < \infty, ... -\infty < X_p < \infty)$$

Además, $E(X_i)$ y $V(X_i)$, para $i = 1, 2, ..., p$, puede determinarse a partir de la distribución de probabilidad marginal de X_i o a partir de la distribución de probabilidad conjunta de $X_1, X_2, ..., X_p$, como se indica a continuación.

$$E(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} x_i f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) dx_1\, dx_2 ... dx_p$$

y (6-23)

$$V(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} (x_i - \mu_{X_i})^2 f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) dx_1\, dx_2 ... dx_p$$

La distribución de probabilidad de un subconjunto de variables como $X_1, X_2, ..., X_k$, $k < p$, puede obtenerse a partir de la distribución de probabilidad conjunta de $X_1, X_2, X_3, ..., X_p$ de la siguiente manera:

Si la función de densidad de probabilidad conjunta de las variables aleatorias continuas $X_1, X_2, ..., X_p$ es $f_{X_1X_2...X_p}(x_1, x_2, ..., x_p)$, entonces la **función de densidad de probabilidad** de $X_1, X_2, ..., X_k$, $k < p$, es

$$= \int \int_{R_{x_1x_2...x_k}} ... \int f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) dx_{k+1} dx_{k+2} ... dx_p \qquad (6\text{-}24)$$

donde $R_{x_1x_2...x_k}$ denota el conjunto de todos los puntos en el rango $X_1, X_2, ..., X_k$ para los que $X_1 = x_1, X_2 = x_2, ..., X_k = x_k$.

Asimismo, pueden desarrollarse distribuciones de probabilidad condicional para variables aleatorias continuas múltiples, extendiendo las ideas usadas para dos variables aleatorias continuas. Por ejemplo, la función de densidad de probabilidad conjunta condicional de X_1, X_2, X_3, dadas X_4, X_5, es

$$f_{X_1X_2X_3|x_4x_5}(x_1, x_2, x_3) = f_{X_1X_2X_3X_4X_5}(x_1, x_2, x_3, x_4, x_5) / f_{X_4X_5}(x_4, x_5)$$

para $f_{X_4X_5}(x_4, x_5) > 0$.

El concepto de independencia puede extenderse a variables aleatorias continuas múltiples.

Definición

Las variables aleatorias continuas $X_1, X_2, ..., X_p$ son **independientes** si y sólo si

$$f_{X_1X_2...X_p}(x_1 x_2, ..., x_p) = f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_p}(x_p) \quad \text{para } toda\ x_1, x_2, ..., x_p \qquad (6\text{-}25)$$

En forma similar al resultado para únicamente dos variables aleatorias, la independencia implica que la ecuación 6-25 es válida para *toda* $x_1, x_2, ..., x_p$. Si se encuentra un punto para el que la igualdad no sea válida, entonces $X_1, X_2, ..., X_p$ no son independientes. Se deja como ejercicio demostrar que si $X_1, X_2, ..., X_p$ son independientes, entonces

$$P(X_1 \in A_1, X_2 \in A_2, ..., X_p \in A_p) = P(X_1 \in A_1) P(X_2 \in A_2) ... P(X_p \in A_p)$$

para cualesquiera regiones $A_1, A_2, ..., A_p$ en el rango de $X_1, X_2, ..., X_p$, respectivamente.

EJEMPLO 6-24

En el capítulo 4 se demostró que una variable aleatoria binomial negativa con parámetros p y r puede representarse como una suma de r variables aleatorias geométricas $X_1, X_2, ..., X_r$. Cada variable aleatoria geométrica representa los ensayos adicionales requeridos para obtener el siguiente éxito. Puesto que los ensayos en un experimento binomial son independientes, $X_1, X_2, ..., X_r$ son variables aleatorias independientes.

EJEMPLO 6-25

Suponga que X_1, X_2 y X_3 representan el espesor en micrones (milésimas de milímetro) de un sustrato, una capa activa y una capa de recubrimiento de un producto químico. Suponga que X_1, X_2 y X_3 son independientes y que tienen una distribución normal con $\mu_1 = 10\,000$, $\mu_2 = 1\,000$, $\mu_3 = 80$, $\sigma_1 = 250$, $\sigma_2 = 20$ y $\sigma_3 = 4$, respectivamente. Las especificaciones para el espesor del sustrato, la capa activa y la capa de recubrimiento son $9\,200 < x_1 < 10\,800$, $950 < x_2 < 1\,050$, y $75 < x_3 < 85$, respectivamente. ¿Qué proporción de los productos químicos cumple con todas las especificaciones? ¿Cuál de los tres espesores es el que tiene la probabilidad menor de cumplir con las especificaciones?

La probabilidad pedida es $P(9\,200 < X_1 < 10\,800, 950 < X_2 < 1\,050, 75 < X_3 < 85)$. Puesto que las variables aleatorias son independientes,

$$\begin{aligned} &P(9\,200 < X_1 < 10\,800\,, 950 < X_2 < 1\,050, 75 < X_3 < 85) \\ &= P(9\,200 < X_1 < 10\,800)\, P(950 < X_2 < 1\,050) P(75 < X_3 < 85) \end{aligned}$$

Después de hacer la estandarización, la expresión anterior es igual a

$$P(-3.2 < Z < 3.2)P(-2.5 < Z < 2.5)P(-1.25 < Z < 1.25)$$

donde Z es una variable aleatoria normal estándar. Por la tabla de la distribución normal estándar, la expresión anterior es igual a

$$(0.99862)(0.98758)(0.78870) = 0.7778$$

El espesor de la capa de recubrimiento es la que tiene la probabilidad menor de cumplir con las especificaciones. Por consiguiente, la prioridad debería ser reducir la variabilidad en esta parte del proceso.

EJERCICIOS DE LA SECCIÓN 6-4

6-55. Suponga que las variables aleatorias X, Y y Z tienen la función de densidad de probabilidad conjunta $f(x, y, z) = 8xyz$ para $0 < x < 1$, $0 < y < 1$, y $0 < z < 1$. Determine lo siguiente:

a) $P(X < 0.5)$

b) $P(X < 0.5, Y < 0.5)$

c) $P(z < 1.5)$

d) $P(X < 0.5$ o $Z < 2)$

e) $E(X)$

6-56. Continuación del ejercicio 6-55. Determine lo siguiente:

a) $P(X < 0.5 \mid Y = 0.5)$

b) $P(X < 0.5, Y < 0.5 \mid Z = 0.8)$

6-57. Continuación del ejercicio 6-55. Determine lo siguiente:
a) La distribución de probabilidad condicional de X, dado que $Y = 0.5$ y $Z = 0.8$.
b) $P(X < 0.5 \mid Y = 0.5, Z = 0.8)$

6-58. Suponga que las variables aleatorias X, Y y Z tienen la función de densidad de probabilidad conjunta $f_{XYZ}(x, y, z) = c$ en el cilindro $x^2 + y^2 < 4$ y que $0 < z < 4$. Determine lo siguiente:
a) La constante c, tal que $f_{XYZ}(x, y, z)$ sea una función de densidad de probabilidad.
b) $P(X^2 + Y^2 < 2)$
c) $P(Z < 2)$
d) $E(X)$

6-59. Continuación del ejercicio 6-58. Determine lo siguiente:
a) $P(X < 1 \mid Y = 1)$
b) $P(X^2 + Y^2 < 1 \mid Z = 1)$

6-60. Continuación del ejercicio 6-58. Determine la distribución de probabilidad condicional de Z, dado que $X = 1$ y $Y = 1$.

6-61. El rendimiento en libras de un día de producción tiene una distribución normal con una media de 1 500 libras y una varianza de 10 000 libras cuadradas. Suponga que el rendimiento en días diferentes son variables aleatorias independientes.
a) ¿Cuál es la probabilidad de que el rendimiento de la producción exceda 1 400 libras en cada uno de los cinco días de la próxima semana?
b) ¿Cuál es la probabilidad de que el rendimiento de la producción exceda 1 400 libras en al menos cuatro de los cinco días de la próxima semana?

6-62. El peso de los ladrillos de adobe usados en la construcción tiene una distribución normal con una media de 3 libras y una desviación estándar de 0.25 libras. Suponga que el peso de los ladrillos es independiente y que se selecciona una muestra aleatoria de 20 ladrillos.
a) ¿Cuál es la probabilidad de que todos los ladrillos de la muestra excedan 2.75 libras?
b) ¿Cuál es la probabilidad de que el ladrillo más pesado de la muestra exceda 3.75 libras?

6-5 COVARIANZA Y CORRELACIÓN

Cuando se definen dos o más variables aleatorias en un espacio de probabilidad, resulta conveniente describir la forma en que varían en conjunto; es decir, es útil medir la relación entre las variables. Una medida común de la relación entre dos variables aleatorias es la **covarianza**. Para definir la covarianza es necesario describir el valor esperado de una función de dos variables aleatorias $h(x, y)$.

Definición

$$E[h(X, Y)] = \begin{cases} \sum\sum_R h(x, y) f_{XY}(x, y) & X, Y \text{ discretas} \\ \iint_R h(x, y) f_{XY}(x, y)\, dx\, dy & X, Y \text{ continuas} \end{cases} \tag{6-26}$$

Es decir, $E[h(X, Y)]$ puede considerarse como el promedio ponderado de $h(x, y)$ para cada punto del rango de (X, Y). El valor de $E[h(X, Y)]$ representa el valor promedio de $h(X, Y)$ que se espera en una serie larga de ensayos repetidos del experimento aleatorio.

EJEMPLO 6-26

Para la distribución de probabilidad conjunta de las dos variables aleatorias de la figura 6-12, calcular $E[(X - \mu_X)(Y - \mu_Y)]$.

El resultado se obtiene multiplicando $x - \mu_X$ por $y - \mu_Y$ por $f_{XY}(x, y)$ para cada punto en el rango de (X, Y). Primero se determinan μ_X y μ_Y como

$$\mu_X = 1\times 0.3 + 3\times 0.7 = 2.4$$

y

$$\mu_Y = 1\times 0.3 + 2\times 0.4 + 3\times 0.3 = 2.0$$

Por lo tanto,

$$\begin{aligned} E[(X-\mu_X)(Y-\mu_Y)] &= (1-2.4)(1-2.0)\times 0.1 \\ &\quad + (1-2.4)(2-2.0)\times 0.2 + (3-2.4)(1-2.0)\times 0.2 \\ &\quad + (3-2.4)(2-2.0)\times 0.2 \\ &\quad + (3-2.4)(3-2.0)\times 0.3 = 0.2 \end{aligned}$$

La covarianza se define para variables aleatorias tanto continuas como discretas por la misma fórmula.

Definición

La **covarianza** entre las variables aleatorias X y Y, denotada como $\text{cov}(X, Y)$ o σ_{XY}, es

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X\mu_Y \qquad (6\text{-}27)$$

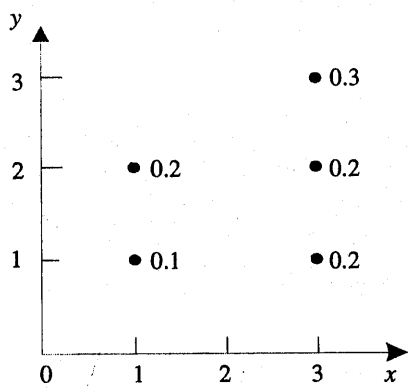


Figura 6-12 Distribución conjunta de X y Y para el ejemplo 6-26.

Si los puntos de la distribución de probabilidad conjunta de X y Y que reciben una probabilidad positiva tienden a caer en una recta con pendiente positiva (o negativa), entonces σ_{XY} es positiva (o negativa). Si los puntos tienden a caer en una recta con pendiente positiva, X tiende a ser mayor que μ_X cuando Y es mayor que μ_Y. Por lo tanto, el producto de los dos términos $x - \mu_X$ y $y - \mu_Y$ tiende a ser positivo. Sin embargo, si los puntos tienden a caer en una recta con pendiente negativa, entonces $x - \mu_X$ tiende a ser positiva cuando $y - \mu_Y$ es negativa, y viceversa. Por lo tanto, el producto de $x - \mu_X$ y $y - \mu_Y$ tiende a ser negativo. En este sentido, la covarianza entre X y Y describe la variación entre las dos variables aleatorias. En la figura 6-13 se presentan ejemplos de pares de variables aleatorias con covarianza positiva, negativa y cero.

La covarianza es una medida de **asociación lineal** entre las variables aleatorias. Si la relación entre las variables aleatorias no es lineal, la covarianza podría no ser sensible a la relación. Esto se ilustra en la figura 6-13*d*. Los únicos puntos con una probabilidad diferente de cero son los puntos en el círculo. Hay una relación identificable entre las variables pero, aun así, la covarianza es cero.

La igualdad de las dos expresiones para la covarianza de la ecuación 6-27 se muestra para variables aleatorias continuas de la siguiente manera: al escribir las expectativas como integrales,

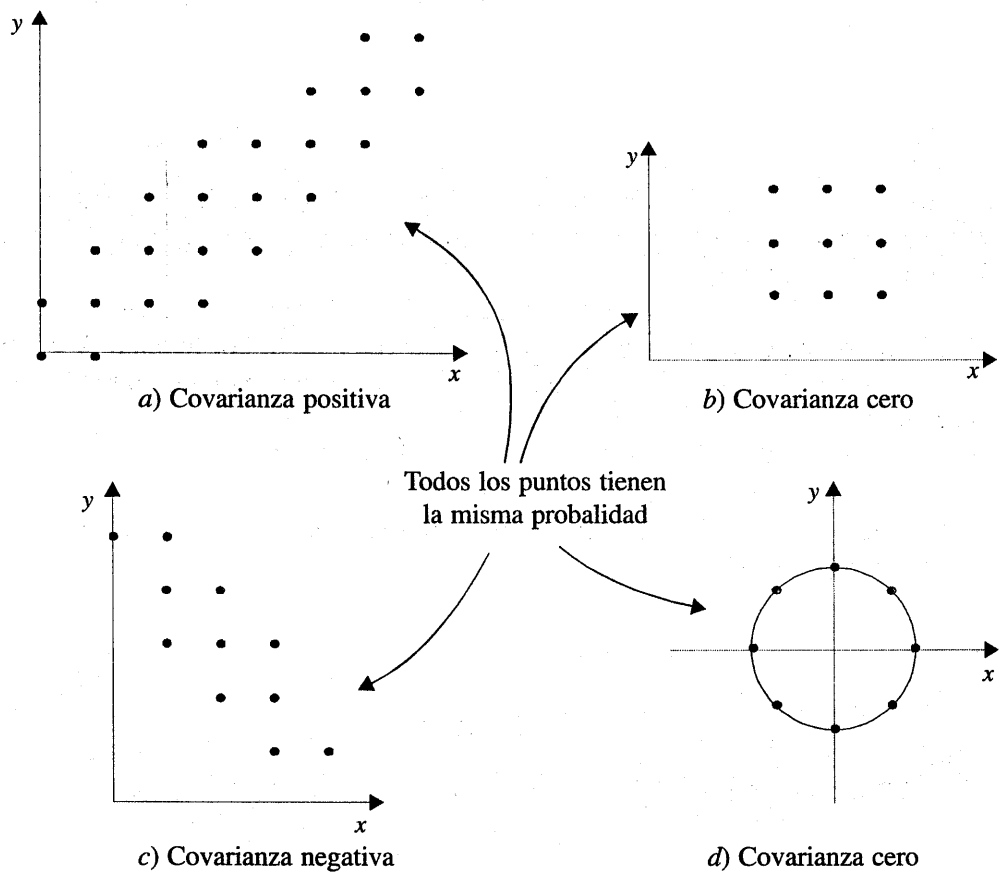


Figura 6-13 Distribuciones de probabilidad conjunta y el signo de la covarianza entre X y Y.

$$E[(Y-\mu_Y)(X-\mu_X)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f_{XY}(x, y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [xy - \mu_X y - x\mu_Y + \mu_X\mu_Y] f_{XY}(x, y)\,dx\,dy$$

Ahora bien,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mu_X y f_{XY}(x, y)\,dx\,dy = \mu_X\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x, y)\,dx\,dy\right] = \mu_X\mu_Y$$

Por lo tanto,

$$E[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)\,dx\,dy - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x, y)\,dx\,dy - \mu_X\mu_Y = E(XY) - \mu_X\mu_Y$$

EJEMPLO 6-27

En el ejemplo 6-1, las variables aleatorias X y Y son el número de bits aceptables y dudosos, respectivamente, entre los cuatro bits recibidos durante una comunicación digital. ¿La covarianza entre X y Y es positiva o negativa?

Puesto que X y Y son el número de bits aceptables y dudosos de los cuatro bits recibidos, $X + Y \leq 4$. Si X está cerca de 4, entonces Y debe estar cerca de 0. Por lo tanto, X y Y tienen una covarianza negativa. Esto puede verificarse a partir de la distribución de probabilidad conjunta de la figura 6-1.

Hay otra medida de la relación entre dos variables aleatorias que muchas veces es más fácil de interpretar que la covarianza.

Definición

La **correlación** entre las variables aleatorias X y Y, denotada como ρ_{XY}, es

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \tag{6-28}$$

Puesto que $\sigma_X > 0$ y $\sigma_Y > 0$, si la covarianza entre X y Y es positiva, negativa o cero, la correlación entre X y Y es positiva, negativa o cero, respectivamente. Es posible demostrar el siguiente resultado.

Para dos variables aleatorias X y Y cualesquiera,

$$-1 \leq \rho_{XY} \leq +1 \tag{6-29}$$

La correlación simplemente escala la covarianza por la desviación estándar de cada variable. Por consiguiente, la correlación es una cantidad adimensional que puede usarse para comparar las relaciones lineales entre pares de variables con unidades diferentes.

Si los puntos de la distribución de probabilidad conjunta de X y Y que reciben una probabilidad positiva tienden a caer en una recta con pendiente positiva (o negativa), entonces ρ_{XY} está cerca de +1 (o de −1). Si ρ_{XY} es igual a +1 o −1, entonces puede demostrarse que los puntos de la distribución de probabilidad conjunta que reciben una probabilidad positiva caen exactamente en una línea recta. Se dice que dos variables aleatorias con correlación diferente de cero están **correlacionadas**. En forma similar a la covarianza, la correlación es una medida de **relación lineal** entre variables aleatorias.

EJEMPLO 6-28

Para las variables aleatorias discretas X y Y con la distribución conjunta que se muestra en la figura 6-14, determine σ_{XY} y ρ_{XY}.

Los cálculos para $E(XY)$, $E(X)$ y $V(X)$ son como sigue:

$$\begin{aligned} E(XY) &= 0 \times 0 \times 0.2 + 1 \times 1 \times 0.1 + 1 \times 2 \times 0.1 + 2 \times 1 \times 0.1 \\ &\quad + 2 \times 2 \times 0.1 + 3 \times 3 \times 0.4 = 4.5 \\ E(X) &= 0 \times 0.2 + 1 \times 0.2 + 2 \times 0.2 + 3 \times 0.4 = 1.8 \\ V(X) &= (0 - 1.8)^2 \times 0.2 + (1 - 1.8)^2 \times 0.2 + (2 - 1.8)^2 \times 0.2 \\ &\quad + (3 - 1.8)^2 \times 0.4 = 1.36 \end{aligned}$$

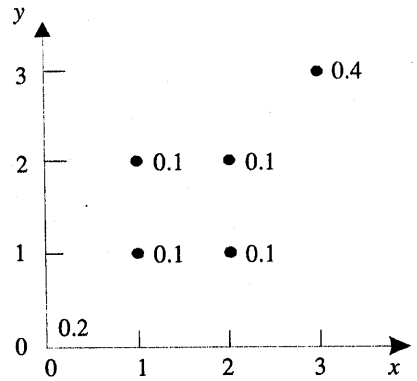


Figura 6-14 Distribución conjunta para el ejemplo 6-28.

Como la distribución de probabilidad marginal de Y es la misma que la de X, $E(Y) = 1.8$ y $V(Y) = 1.36$. Por consiguiente,

$$\sigma_{XY} = E(XY) - E(X)E(Y) = 4.5 - (1.8)(1.8) = 1.26$$

Además,

$$\rho_{XY} = \sigma_{XY}/(\sigma_X \sigma_Y) = 1.26/[(1.36^{1/2})(1.36^{1/2})] = 0.926$$

EJEMPLO 6-29

Suponga que la variable aleatoria X tiene la siguiente distribución: $P(X = 1) = 0.2$, $P(X = 2) = 0.6$, $P(X = 3) = 0.2$. Sea $Y = 2X + 5$. Es decir, $P(Y = 7) = 0.2$, $P(Y = 9) = 0.6$, $P(Y = 11) = 0.2$. Determine la correlación entre X y Y (véase la figura 6-15).

Puesto que X y Y se relacionan linealmente, $\rho = 1$. Esto puede verificarse por medio de cálculos directos. ¡Inténtelo!

En el caso de variables aleatorias independientes, no se espera ninguna relación en su distribución de probabilidad conjunta. El resultado siguiente se deja como ejercicio.

Si X y Y son variables aleatorias independientes, entonces

$$\sigma_{XY} = \rho_{XY} = 0 \tag{6-30}$$

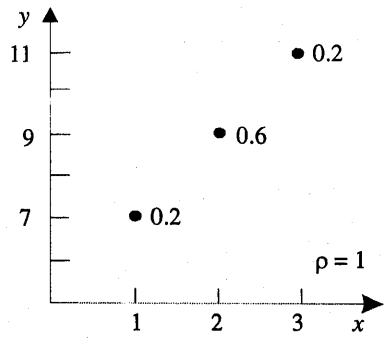


Figura 6-15 Densidad de probabilidad conjunta para el ejemplo 6-29.

EJEMPLO 6-30

Para las dos variables aleatorias de la figura 6-16, demuestre que $\sigma_{XY} = 0$.

Las dos variables aleatorias de este ejemplo son variables aleatorias continuas. En este caso, $E(XY)$ se define como la integral doble en el rango de (X, Y). Es decir,

$$E(XY) = \int_0^4 \int_0^2 xyf_{XY}(x, y)dx\, dy = \frac{1}{16}\int_0^4 \left[\int_0^2 x^2 y^2 dx\right] dy = \frac{1}{16}\int_0^4 y^2\left[x^3/3\Big|_0^2\right]$$

$$= \frac{1}{16}\int_0^4 y^2[8/3]dy = \frac{1}{6}\left[y^3/3\Big|_0^4\right] = \frac{1}{6}[64/3] = 32/9$$

Asimismo,

$$E(X) = \int_0^4 \int_0^2 xf_{XY}(x, y)\, dx\, dy = \frac{1}{16}\int_0^4 \left[\int_0^2 x^2 dx\right] dy = \frac{1}{16}\int_0^4 \left[x^3/3\Big|_0^2\right] dy$$

$$= \frac{1}{16}\left[y^2/2\Big|_0^4\right][8/3] = \frac{1}{6}[16/2] = 4/3$$

y

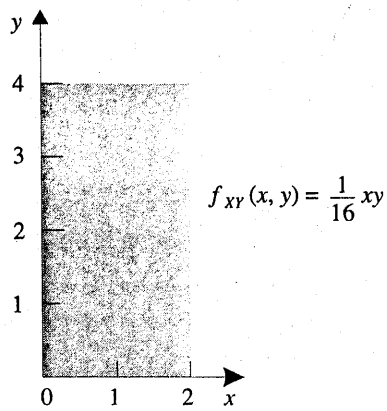


Figura 6-16 Las variables aleatorias con covarianza cero del ejemplo 6-30.

$$E(Y)=\int_0^4\int_0^2 yf_{XY}(x,\ y)\,dx\,dy=\frac{1}{16}\int_0^4 y^2\left[\int_0^2 x\,dx\right]dy=\frac{1}{16}\int_0^4 y^2\left[x^2/2\Big|_0^2\right]dy$$

$$=\frac{2}{16}\left[y^3/3\Big|_0^4\right]=\frac{1}{8}[64/3]=8/3$$

Por tanto,

$$E(XY)-E(X)E(Y)=32/9-(4/3)(8/3)=0$$

Puede demostrarse que estas dos variables aleatorias son independientes. El lector puede comprobar que $f_{XY}(x, y) = f_X(x)f_Y(y)$ para toda x y y.

Sin embargo, si la correlación entre dos variables aleatorias es cero, *no puede* concluirse de inmediato que las variables aleatorias son independientes. Un ejercicio al final de este capítulo proporciona un ejemplo.

EJERCICIOS DE LA SECCIÓN 6-5

6-63. Determine la covarianza y la correlación para la siguiente distribución de probabilidad conjunta:

x	1	1	2	3
y	3	4	5	6
$f_{XY}(x, y)$	1/8	1/4	1/2	1/8

6-64. Determine la covarianza y la correlación para la siguiente distribución de probabilidad conjunta:

x	–1	–0.5	0.5	1
y	–2	–1	1	2
$f_{XY}(x, y)$	1/8	1/4	1/2	1/8

6-65. Determine el valor de c, la covarianza y la correlación para la función de masa de probabilidad conjunta $f_{XY}(x, y) = c(x + y)$ para $x = 1, 2, 3$ y $y = 1, 2, 3$.

6-66. Determine la covarianza y la correlación para la distribución de probabilidad conjunta que se muestra en la figura 6-4*a* y que se describe en el ejemplo 6-8.

6-67. Determine la covarianza y la correlación para la distribución de probabilidad conjunta de las variables aleatorias X_2 y X_3 del ejemplo 6-13. ¿A qué conclusión puede llegarse respecto del signo de la correlación entre dos variables aleatorias de una distribución multinomial?

6-68. Determine el valor de c, la covarianza y la correlación para la función de densidad de probabilidad conjunta $f_{XY}(x, y) = cxy$ en el rango $0 < x < 3$ y $0 < y < x$.

6-69. Determine el valor de c, la covarianza y la correlación para la función de densidad de probabilidad conjunta $f_{XY}(x, y) = c$ en el rango $0 < x < 4$, $0 < y$ y $x - 1 < y < x + 1$.

6-70. Determine la covarianza y la correlación para la función de densidad de probabilidad conjunta $f_{XY}(x, y) = 6 \times 10^{-6}e^{-0.001x - 0.002y}$ en el rango $0 < x$ y $x < y$ del ejemplo 6-14.

6-71. Determine la covarianza y la correlación para la función de densidad de probabilidad conjunta $f_{XY}(x, y) = e^{-x-y}$ en el rango $0 < x$ y $0 < y$. La distribución de probabilidad conjunta es:

x	−1	0	0	1
y	0	−1	1	0
$f_{XY}(x, y)$	1/4	1/4	1/4	1/4

6-72. Demuestre que la correlación entre X y Y es cero, pero X y Y no son independientes. Suponga que la correlación entre X y Y es ρ. Para las constantes a, b, c y d, ¿cuál es la correlación entre las variables aleatorias $U = aX + b$ y $V = cY + d$?

6-73. Suponga que X y Y son variables aleatorias continuas independientes. Demuestre que $\sigma_{XY} = 0$.

6-6 DISTRIBUCIÓN NORMAL DE DOS VARIABLES

La extensión de la distribución normal para dos variables aleatorias es una importante distribución de probabilidad bivariada.

EJEMPLO 6-31

Al principio de este capítulo se presentó la longitud de diferentes dimensiones de una pieza moldeada por inyección como un ejemplo de dos variables aleatorias. Cada longitud podría modelarse con una distribución normal. Sin embargo, debido a que las mediciones son de la misma pieza, las variables aleatorias por lo general no son independientes. Una distribución de probabilidad para dos variables aleatorias normales que no son independientes es importante en muchas aplicaciones. Como se estableció al principio de este capítulo, si las especificaciones de X y Y son (2.95 a 3.05) y (7.60 a 7.80) milímetros, respectivamente, entonces podría interesarnos la probabilidad de que una pieza satisfaga ambas especificaciones; es decir, $P(2.95 < X < 3.05, 7.60 < Y < 7.80)$.

Definición

La función de densidad de probabilidad de una distribución normal de dos variables es

$$f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right\} \tag{6-31}$$

para $-\infty < x < \infty$ y $-\infty < y < \infty$, con parámetros $\sigma_X > 0$, $\sigma_Y > 0$, $-\infty < \mu_X < \infty$, $-\infty < \mu_Y < \infty$, y $-1 < \rho < 1$.

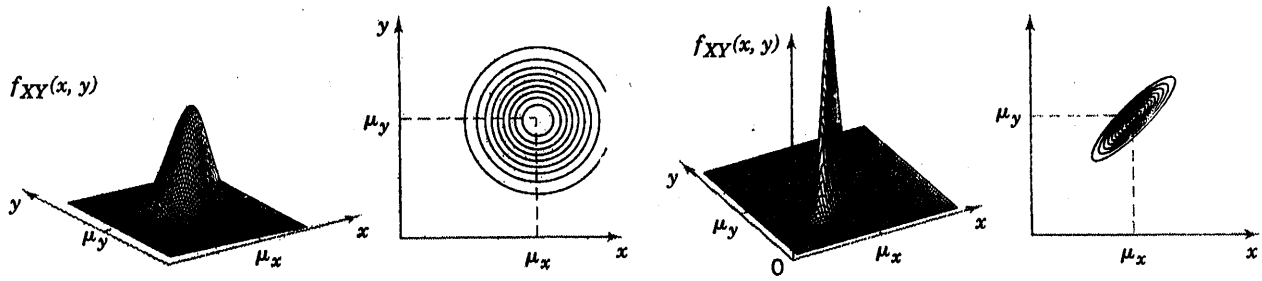


Figura 6-17 Ejemplos de distribuciones normales de dos variables.

El resultado de que la integral de $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$ es uno se deja como ejercicio. Asimismo, la función de densidad de probabilidad normal de dos variables es positiva en todo el plano de los números reales.

En la figura 6-17 se presentan dos ejemplos de distribuciones normales de dos variables junto con las gráficas de contorno correspondientes. Cada curva de las gráficas de contorno es un conjunto de puntos para los que la función de densidad de probabilidad es constante. Como puede verse en las gráficas de contorno, la función de densidad de probabilidad normal de dos variables es constante en elipses en el plano (x, y). (Una circunferencia puede considerarse un caso especial de una elipse.) El centro de cada elipse está en el punto (μ_X, μ_Y). Si $\rho > 0$ $(\rho < 0)$, el eje mayor de cada elipse tiene pendiente positiva (negativa), respectivamente. Si $\rho = 0$, el eje mayor de la elipse es paralelo al eje de coordenadas x o bien al eje y.

EJEMPLO 6-32

La función de densidad de probabilidad conjunta $f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-0.5(x^2+y^2)}$ es un caso especial de una distribución normal de dos variables con $\sigma_X = 1$, $\sigma_Y = 1$, $\mu_X = 0$, $\mu_Y = 0$ y $\rho = 0$. Esta función de densidad de probabilidad se ilustra en la figura 6-18. Obsérvese que la gráfica de contorno consiste en círculos concéntricos alrededor del origen.

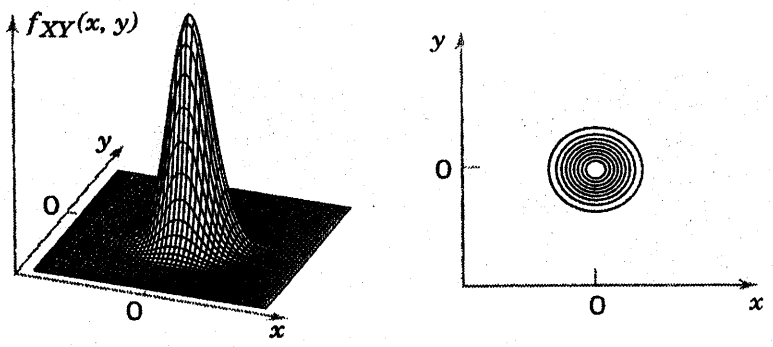


Figura 6-18 Función de densidad de probabilidad normal de dos variables con $\sigma_X = 1$, $\sigma_Y = 1$, $\mu_X = 0$ y $\mu_Y = 0$.

Al completar el trinomio cuadrado perfecto en el exponente, es posible demostrar los siguientes resultados. Los detalles se dejan como ejercicio.

Si X y Y tienen una distribución normal de dos variables con densidad de probabilidad conjunta $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$, entonces las *distribuciones de probabilidad marginal* de X y Y son normales con medias μ_X y μ_Y y desviaciones estándar σ_X y σ_Y, respectivamente. (6-32)

En la figura 6-19 se ilustra que las distribuciones de probabilidad marginal de X y Y son normales. Además, como lo sugiere la notación, r representa la correlación entre X y Y. El siguiente resultado se deja como ejercicio.

Si X y Y tienen una distribución normal de dos variables con densidad de probabilidad conjunta $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$, entonces la correlación entre X y Y es ρ. (6-33)

Las gráficas de contorno de la figura 6-17 ilustran que cuando r pasa de cero (gráfica izquierda) a 0.9 (gráfica derecha), las elipses se angostan alrededor del eje mayor. La probabilidad se concentra más alrededor de una recta en el plano (x, y) y muestra gráficamente una correlación mayor entre las variables. Si $\rho = -1$ o $+1$, entonces toda la probabilidad se concentra en una recta en el plano (x, y). Es decir, es cero la probabilidad de que X y Y asuman un valor que no esté sobre la recta. En este caso, la densidad de probabilidad normal de dos variables no está definida.

En el caso especial en que X y Y tienen una distribución normal de dos variables, si $\rho = 0$, entonces X y Y son independientes. Los detalles se dejan como ejercicio.

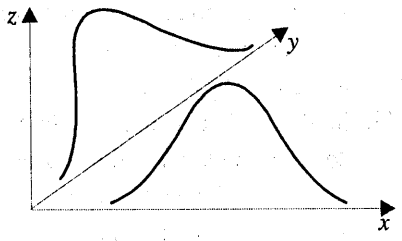


Figura 6-19 Funciones de densidad de probabilidad marginal de una distribución normal de dos variables.

Si X y Y tienen una distribución normal de dos variables con $\rho = 0$, entonces X y Y son independientes. (6-34)

Una aplicación importante de la distribución normal de dos variables es para calcular probabilidades cuando intervienen dos variables aleatorias normales correlacionadas.

EJEMPLO 6-33

Suponga que las dimensiones X y Y de una pieza moldeada por inyección tienen una distribución normal de dos variables con $\sigma_X = 0.04$, $\sigma_Y = 0.08$, $\mu_X = 3.00$, $\mu_Y = 7.70$ y $\rho = 0.8$. Entonces, la probabilidad de que una pieza satisfaga ambas especificaciones es

$$P(2.95 < X < 3.05, 7.60 < Y < 7.80)$$

Esta probabilidad puede obtenerse integrando $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$ en la región $2.95 < x < 3.05$ y $7.60 < y < 7.80$, como se muestra en la figura 6-7. Desafortunadamente, en la mayoría de los casos no se cuenta con una solución de forma cerrada para las probabilidades en las que intervienen distribuciones normales de dos variables. En este caso, la integración debe hacerse por métodos numéricos.

EJERCICIOS DE LA SECCIÓN 6-6

6-74. Sea que X y Y representen la concentración y la viscosidad de un producto químico. Suponga que X y Y tienen una distribución normal de dos variables con $\sigma_X = 4$, $\sigma_Y = 1$, $\mu_X = 2$ y $\mu_Y = 1$. Trace una gráfica de contorno aproximada de la función de densidad de probabilidad conjunta para cada uno de los siguientes valores de ρ:

a) $\rho = 0$

b) $\rho = 0.8$

c) $\rho = -0.8$

6-75. Sea que X y Y representan dos dimensiones de una pieza moldeada por inyección. Suponga que X y Y tienen una distribución normal de dos variables con $\sigma_X = 0.04$, $\sigma_Y = 0.08$, $\mu_X = 3.00$, $\mu_Y = 7.70$ y $\rho = 0$. Determine $P(2.95 < X < 3.05, 7.60 < Y < 7.80)$.

6-76. Demuestre que la integral de la función de densidad de probabilidad $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$ de una distribución normal de dos variables es uno. [*Sugerencia:* Complete el trinomio cuadrado perfecto en el exponente y use el hecho de que la integral de una función de densidad de probabilidad normal para una sola variable es uno.]

6-77. Si X y Y tienen una distribución normal de dos variables con densidad de probabilidad conjunta $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$, demuestre que la *distribución de probabilidad marginal* de X es normal con media μ_X y desviación estándar σ_X. [*Sugerencia:* Complete el trinomio cuadrado perfecto en el exponente y use el hecho de que la integral de una función de densidad de probabilidad normal para una sola variable es uno.]

6-78. Si X y Y tienen una distribución normal de dos variables con densidad de probabilidad conjunta $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$, demuestre que la correlación entre X y Y es ρ. [*Sugerencia:* Complete el trinomio cuadrado perfecto en el exponente.]

6-7 COMBINACIONES LINEALES DE VARIABLES ALEATORIAS

Una variable aleatoria se define en ocasiones como una función de varias variables aleatorias. Por ejemplo, si las variables aleatorias X_1 y X_2 denotan la longitud y el ancho, respectivamente, de una pieza manufacturada, entonces $Y = 2X_1 + 2X_2$ es una variable aleatoria que representa el perímetro de la pieza. Como otro ejemplo, recuérdese que la variable aleatoria binomial negativa se representó como la suma de varias variables aleatorias geométricas.

En esta sección se desarrollan resultados para variables aleatorias que son combinaciones lineales de variables aleatorias.

Definición

Dadas las variables aleatorias $X_1, X_2, \ldots, X_p$ y las constantes $c_1, c_2, \ldots, c_p$, entonces

$$Y = c_1X_1 + c_2X_2 + \ldots + c_pX_p \tag{6-35}$$

es una **combinación lineal** de $X_1, X_2, \ldots, X_p$.

Ahora bien, $E(Y)$ puede encontrarse a partir de la distribución de probabilidad conjunta de X_1, X_2, ..., X_p como se indica a continuación. Suponga que X_1, X_2, ..., X_p son variables aleatorias continuas. Puede aplicarse un cálculo similar para variables aleatorias discretas.

$$\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (c_1x_1 + c_2x_2 + \cdots + c_px_p) f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1dx_2\ldots dx_p \\
&= c_1\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_1 f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1dx_2\ldots dx_p \\
&\quad + c_2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_2 f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1dx_2\ldots dx_p \\
&\quad +, \cdots, \\
&\quad + c_p\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_p f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\,dx_1dx_2\ldots dx_p
\end{aligned}$$

Al aplicar la ecuación 6-23 en cada término de esta expresión, se llega al siguiente resultado:

Si $Y = c_1X_1 + c_2X_2 + \ldots + c_pX_p$, entonces

$$E(Y) = c_1E(X_1) + c_2E(X_2) + \ldots + c_pE(X_p) \qquad (6\text{-}36)$$

Además, se deja como ejercicio demostrar el siguiente resultado.

Si $X_1, X_2, \ldots, X_p$ son variables aleatorias, y $Y = c_1X_1 + c_2X_2 + \ldots + c_pX_p$, entonces en general

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \ldots + c_p^2V(X_p) \qquad 6\text{-}37)$$

$$+2\sum\sum_{i<j=2}^{p} c_ic_j\,\mathrm{cov}(X_i, X_j)$$

Además, si $X_1, X_2, \ldots, X_p$ son **independientes**, entonces

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \ldots + c_p^2V(X_p) \qquad (6\text{-}38)$$

Obsérvese que el resultado para la varianza requiere que las variables aleatorias sean independientes. Para ver porqué es importante la independencia, considérese el siguiente ejemplo simple. Sea X_1 cualquier variable aleatoria y se define $X_2 = -X_1$. Evidentemente, X_1 y X_2 no son independientes. De hecho, $\rho_{XY} = -1$. Ahora bien, $Y = X_1 + X_2$ es 0, con probabilidad uno. Por lo tanto, $V(Y) = 0$, independientemente de las varianzas de X_1 y X_2.

EJEMPLO 6-34

Los resultados de las ecuaciones 6-36 y 6-37 pueden usarse para obtener la media y la varianza de una variable aleatoria binomial negativa.

En el capítulo 4 se encontró que si Y es una variable aleatoria binomial negativa con parámetros p y r, entonces $Y = X_1 + X_2 + \ldots + X_r$, donde cada X_i es una variable aleatoria geométrica con parámetro p. Por lo tanto, $E(X_i) = 1/p$ y $V(X_i) = (1 - p)/p^2$. Por la ecuación 6-36, $E(Y) = r/p$. Además, en el capítulo 4 se estableció que $X_1, X_2, \ldots, X_p$ son independientes. Por lo tanto, por la ecuación 6-38, $V(Y) = r(1 - p)/p^2$.

Puede usarse un enfoque similar al que se aplicó en el ejemplo anterior para verificar las fórmulas de la media y la varianza de una variable aleatoria de Erlang del capítulo 5. En seguida, considérese otra combinación lineal.

EJEMPLO 6-35

Suponga que las variables aleatorias X_1 y X_2 denotan la longitud y el ancho, respectivamente, de una pieza manufacturada. Suponga que $E(X_1) = 2$ centímetros con desviación estándar de 0.1 centímetros y que $E(X_2) = 5$ centímetros con desviación estándar de 0.2 centímetros. Asimismo, suponga que X_1 y X_2 son independientes. Entonces, $Y = 2X_1 + 2Y_2$ es una variable aleatoria que representa el perímetro de la pieza. Por la ecuación 6-36,

$$E(Y) = 2(2) + 2(5) = 14 \text{ cm}$$

y por la ecuación 6-38,

$$V(Y) = 2^2(0.1^2) + 2^2(0.2^2) = 0.04 + 0.16 = 0.20 \text{ cm}^2$$

Por lo tanto, la desviación estándar de Y es $0.20^{1/2} = 0.447$ cm.

La combinación lineal particular que representa el promedio de p variables aleatorias, con medias y varianzas idénticas, se usa con mucha frecuencia en los capítulos subsecuentes de este libro. Se destacan los resultados para este caso especial.

Si $\overline{X} = (X_1 + X_2 + \cdots + X_p)/p$ con $E(X_i) = \mu$ para $i = 1, 2, \ldots, p$, entonces

$$E(\overline{X}) = \mu$$

Además, si $X_1, X_2, \ldots, X_p$ también son **independientes** con $V(X_i) = \sigma^2$ para $i = 1, 2, \ldots, p$, entonces

$$V(\overline{X}) = \sigma^2/p \tag{6-39}$$

La conclusión para $V(\overline{X})$ se obtiene de la siguiente manera. Utilizando la ecuación 6-38, con $c_i = 1/p$ y $V(X_i) = \sigma^2$, se obtiene

$$V(\overline{X}) = \underbrace{(1/p)^2\sigma^2 + \cdots + (1/p)^2\sigma^2}_{p \text{ términos}} = \sigma^2/p$$

Otro resultado útil respecto a las combinaciones lineales de variables aleatorias es una propiedad reproductiva que es válida para variables aleatorias normales e independientes.

Propiedad reproductiva de la distribución normal

Si $X_1, X_2, ..., X_p$ son variables aleatorias **normales e independientes** con $E(X_i) = \mu_i$ y $V(X_i) = \sigma_i^2$, para $i = 1, 2, ..., p$, entonces

$$Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$$

es una variable aleatoria **normal** con

$$E(Y) = c_1\mu_1 + c_2\mu_2 + \cdots + c_p\mu_p$$

y

$$V(Y) = c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \cdots + c_p^2\sigma_p^2 \qquad (6\text{-}40)$$

La media y la varianza de Y se siguen de las ecuaciones 6-36 y 6-38. El hecho de que Y tiene una distribución normal puede obtenerse a partir de las funciones generadoras que se exponen en el apéndice II.

EJEMPLO 6-36

En el ejemplo 6-35, las variables aleatorias X_1 y X_2 denotan la longitud y el ancho, respectivamente, de una pieza manufacturada. Suponga que X_1 es normal con $E(X_1) = 2$ centímetros y desviación estándar de 0.1 centímetros y que X_2 es normal con $E(X_2) = 5$ centímetros y una desviación estándar de 0.2 centímetros. Suponga asimismo que X_1 y X_2 son independientes. Determine la probabilidad de que el perímetro exceda 14.5 centímetros.

Así, $Y = 2X_1 + 2Y_2$ es una variable aleatoria normal que representa el perímetro de la pieza. Por el ejemplo 6-35, $E(Y) = 14$ centímetros y la desviación estándar de Y es 0.447 cm. Entonces,

$$P(Y > 14.5) = P[(Y - \mu_Y)/\sigma_Y > (14.5 - 14)/0.447]$$
$$= P(Z > 1.12) = 0.13$$

EJEMPLO 6-37

Se llenan latas de refresco con una máquina de llenado automático. El volumen de llenado promedio es 12.1 onzas líquidas, y la desviación estándar es 0.05 onzas líquidas. Suponga que el volumen de llenado de las latas es una variable aleatoria normal e independiente. ¿Cuál es la probabilidad de que el volumen promedio de 10 latas seleccionadas de este proceso sea menor que 12 onzas líquidas?

Sea que $X_1, X_2, ..., X_{10}$ denoten los volúmenes de llenado de las 10 latas. El volumen de llenado promedio (denotado como $\overline{X}$) es una variable aleatoria normal con

$$E(\bar{X}) = 12.1 \quad \text{y} \quad V(\bar{X}) = \frac{0.05^2}{10}\,\text{cm}^2 = 0.00025$$

Por consiguiente,

$$P(\bar{X} < 12) = P\left[(\bar{X} - \mu_{\bar{X}})/\sigma_{\bar{X}} < (12 - 12.1)/\sqrt{0.00025}\right]$$
$$= P(Z < -6.32) = 0$$

EJERCICIOS DE LA SECCIÓN 6-7

6-79. Si X y Y son variables aleatorias normales e independientes con $E(X) = 0$, $V(X) = 4$, $E(Y) = 10$, y $V(Y) = 9$, determine lo siguiente:
a) $E(2X + 3Y)$
b) $V(2X + 3Y)$
c) $P(2X + 3Y < 30)$
d) $P(2X + 3Y < 40)$

6-80. Suponga que la variable aleatoria X representa la longitud en pulgadas de una pieza troquelada. Sea Y la longitud de la pieza en milímetros. Si $E(X) = 5$ pulgadas y $V(X) = 0.25$, ¿cuál es la media y la varianza de Y?

6-81. Un estuche de plástico para discos magnéticos se compone de mitades. El grosor de cada mitad tiene una distribución normal con una media de 1.5 milímetros y una desviación estándar de 0.1 milímetros, y las mitades son independientes.
a) Determine la media y la desviación estándar del grosor total de las dos mitades.
b) ¿Cuál es la probabilidad de que el grosor total exceda 3.3 milímetros?

6-82. La anchura del marco de una puerta tiene una distribución normal con una media de 24 pulgadas y una desviación estándar de 1/8 de pulgada. La anchura de la puerta tiene una distribución normal con una media de 23 y 7/8 de pulgada y una desviación estándar de 1/16 de pulgada. Suponga la independencia.
a) Determine la media y la desviación estándar de la diferencia entre la anchura del marco y la anchura de la puerta.
b) ¿Cuál es la probabilidad de que la anchura del marco menos la anchura de la puerta exceda 1/4 de pulgada?
c) ¿Cuál es la probabilidad de que la puerta no coincida con el marco?

6-83. Un componente en forma de "U" va a formarse con las tres partes A, B y C. En la figura 6-20 se muestra un diagrama. La longitud de A tiene una distribución normal con una media de 10 milímetros y una desviación estándar de 0.1 milímetros. El grosor de las partes B y C tiene una distribución normal con una media de 2 milímetros y una desviación estándar de 0.05 milímetros. Suponga que todas las dimensiones son independientes.
a) Determine la media y la desviación estándar de la longitud de la abertura D.
b) ¿Cuál es la probabilidad de que la abertura D sea menor que 5.9 milímetros?

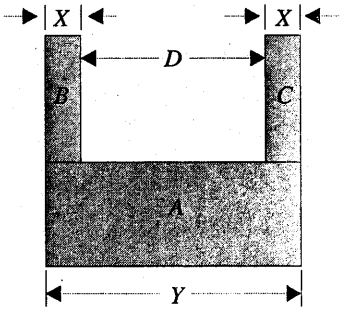


Figura 6-20 Diagrama del componente en forma de "U".

6-84. Se llenan latas de refresco con una máquina de llenado automático, y la desviación estándar es 0.5 onzas líquidas. Suponga que el volumen de llenado de las latas es una variable aleatoria normal e independiente.

a) ¿Cuál es la desviación estándar del volumen de llenado promedio de 100 latas?

b) Si la media del volumen de llenado es 12.1 onzas, ¿cuál es la probabilidad de que el volumen de llenado promedio de las 100 latas esté abajo de 12 onzas líquidas?

c) ¿A qué debería ser igual la media del volumen de llenado para que la probabilidad sea 0.005, de que el promedio de 100 latas esté abajo de 12 onzas líquidas?

6-85. El espesor del fotoprotector en la fabricación de semiconductores tiene una media de 10 micrones (milésimas de milímetro) y una desviación estándar de 1 micrón (milésima de milímetro). Suponga que el espesor tiene una distribución normal y que el espesor de las diferentes obleas es independiente.

a) Determine la probabilidad de que el espesor promedio de 10 obleas sea mayor que 11 o bien menor que 9 micrones (milésimas de milímetro).

b) Determine el número de obleas que es necesario medir para que sea 0.01 la probabilidad de que el espesor promedio exceda 11 micrones (milésimas de milímetro).

6-86. Suponga que el peso de individuos es independiente y que tiene una distribución normal con una media de 160 libras y una desviación estándar de 30 libras. Suponga que se apiñan 25 personas en un elevador que está diseñado para soportar 4 500 libras.

a) ¿Cuál es la probabilidad de que la carga (el peso total) exceda el límite de diseño?

b) ¿Cuál es el número máximo de ocupantes tal que la probabilidad de que el límite de diseño se exceda sea 0.0001? [*Sugerencia:* Un enfoque iterativo podría resultar más sencillo.]

6-8 DESIGUALDAD DE CHEBYSHEV

En el capítulo 4 se indicó que si X es una variable aleatoria normal con media μ y desviación estándar σ, entonces $P(\mu - 1.96\sigma < X < \mu + 1.96\sigma) = 0.95$. Este resultado relaciona la probabilidad de una variable aleatoria normal con la magnitud de la desviación estándar. Un interesante resultado similar que se aplica a cualquier variable aleatoria continua o discreta fue desarrollado por el matemático Chebyshev en 1867.

Desigualdad de Chebyshev

Para cualquier variable aleatoria X con media μ y varianza σ^2,

$$P(|X - \mu| \geq c\sigma) \leq 1/c^2$$

para $c > 0$.

Este resultado tiene la siguiente interpretación. La probabilidad de que una variable aleatoria difiera de su media en al menos c desviaciones estándar es menor o igual que $1/c^2$. Obsérvese que la regla sólo es útil para $c > 1$.

Por ejemplo, usar $c = 2$ implica que la probabilidad de que *cualquier* variable aleatoria difiera de su media en al menos dos desviaciones estándar no es mayor que 1/4. Sabemos que para una variable aleatoria normal esta probabilidad es menor que 0.05. Asimismo, usar $c = 3$ implica que la probabilidad de que cualquier variable aleatoria difiera de su media en al menos tres desviaciones estándar no es mayor que 1/9. La desigualdad de Chebyshev proporciona una relación entre la desviación estándar y la dispersión de la distribución de probabilidad de cualquier variable aleatoria. La demostración se deja como ejercicio.

En la tabla 6-1 se hace la comparación de las probabilidades calculadas con la regla de Chebyshev y las probabilidades calculadas para una variable aleatoria normal.

Tabla 6-1 Porcentaje de una distribución mayor que c desviaciones estándar de la media

c	Regla de Chebyshev para cualquier distribución de probabilidad	Distribución normal
1.5	menor que 44.4%	13.4%
2	menor que 25.0%	4.6%
3	menor que 11.1%	0.27%
4	menor que 6.3%	0.01%

EJEMPLO 6-38

El proceso de taladrado de agujeros en tarjetas de circuitos impresos produce diámetros con una desviación estándar de 0.01 milímetros. ¿Cuántos diámetros es necesario medir para que sea al menos 8/9 la probabilidad de que el promedio de los diámetros medidos esté dentro de 0.005 milímetros del diámetro medio del proceso μ?

Sean $X_1, X_2, \ldots, X_n$ las variables aleatorias que denotan los diámetros de n agujeros. El diámetro promedio medido es $\overline{X} = (X_1 + X_2 + \ldots + X_n)/n$. Suponga que las X son variables aleatorias independientes. Por la ecuación 6-38, $E(\overline{X}) = \mu$ y $V(\overline{X}) = 0.01^2/n$. Por consiguiente, la desviación estándar de $\overline{X}$ es $(0.01^2/n)^{1/2}$. Al aplicar la desigualdad de Chebyshev a $\overline{X}$,

$$P\left(|\overline{X} - \mu| \geq c(0.01^2/n)^{1/2}\right) \leq 1/c^2$$

Sea $c = 3$. Entonces,

$$P\left(|\overline{X} - \mu| \geq 3(0.01^2/n)^{1/2}\right) \leq 1/9$$

Por lo tanto,

$$P\left(|\overline{X} - \mu| < 3(0.01^2/n)^{1/2}\right) \leq 8/9$$

Así, la probabilidad de que $\overline{X}$ esté dentro de $3(0.01^2/n)^{1/2}$ de μ es al menos de 8/9. Por último, n se elige de tal modo que $3(0.01^2/n)^{1/2} = 0.005$. Es decir,

$$n = 3^2\left[0.01^2 / 0.005^2\right] = 36$$

EJERCICIOS DE LA SECCIÓN 6-8

6-87. El espesor del fotoprotector en la fabricación de semiconductores tiene una media de 10 micrones (milésimas de milímetro) y una desviación estándar de 1 micrón (milésima de milímetro). Acote la probabilidad de que el espesor sea menor que 6 o mayor que 14 micrones (milésimas de milímetro).

6-88. Suponga que X tiene una distribución uniforme continua con rango $0 < x < 10$. Use la regla de Chebyshev para acotar la probabilidad de que X difiera en más de dos desviaciones estándar de su media y haga la comparación con la probabilidad real.

6-89. Suponga que X tiene una distribución exponencial con media 20. Use la regla de Chebyshev para acotar la probabilidad de que X difiera de su media en más de dos desviaciones estándar y en más de tres desviaciones estándar, y haga la comparación con las probabilidades reales.

6-90. Suponga que X tiene una distribución de Poisson con media $\lambda = 4$. Use la regla de Chebyshev para acotar la probabilidad de que X difiera de su media en más de dos desviaciones estándar y en más de tres desviaciones estándar, y haga la comparación con las probabilidades reales.

6-91. Considere el proceso de taladrado de agujeros en tablas de circuitos impresos. Suponga que la desviación estándar de los diámetros es 0.01 y que los diámetros son independientes. Suponga que se usa el promedio de 500 diámetros para estimar la media del proceso.

a) La probabilidad de que el promedio medido esté dentro de una cota de la media del proceso, es de al menos 15/16. ¿Cuál es la cota?

b) Si se supone que los diámetros tienen una distribución normal, determine la cota tal que sea 15/16 la probabilidad de que el promedio medido esté más cerca de la media del proceso que la cota.

Ejercicios complementarios

6-92. Demuestre que la siguiente función satisface las propiedades de una función de masa de probabilidad conjunta:

x	y	$f(x, y)$
0	1	1/8
1	0	1/8
1	1	1/4
2	2	1/2

6-93. Continuación del ejercicio 6-92. Determine las siguientes probabilidades:

a) $P(X < 0.5, Y < 1.5)$

b) $P(X \leq 1)$

c) $P(Y < 1.5)$

d) $P(X > 0.5, Y < 1.5)$

6-94. Continuación del ejercicio 6-92. Determine $E(X)$, $E(Y)$, $V(X)$ y $V(Y)$.

6-95. Continuación del ejercicio 6-92. Determine lo siguiente:

a) La distribución de probabilidad marginal de la variable aleatoria X.

b) La distribución de probabilidad condicional de Y, dado que $X = 1$.

c) $E(Y \mid x = 1)$

6-96. Continuación del ejercicio 6-92. ¿Son independientes X y Y? ¿Por qué sí o por qué no?

6-97. Continuación del ejercicio 6-92. Calcule la correlación entre X y Y.

6-98. El torque de remoción requerido para quitar los pernos de una placa de acero se clasifica como alto, moderado y bajo. Históricamente, la probabilidad de una clasificación del torque como alto, moderado o bajo es 0.6, 0.3 o 0.1, respectivamente. Suponga que se evalúan 20 pernos y que las clasificaciones del torque son independientes.

a) ¿Cuál es la probabilidad de que 12, 6 y 2 pernos se clasifiquen como altos, moderados y bajos, respectivamente?

b) ¿Cuál es la distribución marginal del número de pernos clasificados como bajos?

c) ¿Cuál es el número esperado de pernos clasificados como bajos?

d) ¿Cuál es la probabilidad de que el número de pernos clasificados como bajos sea mayor que dos?

6-99. Continuación del ejercicio 6-97.

a) ¿Cuál es la distribución condicional del número de pernos clasificados como bajos, dado que 16 pernos se clasifican como altos?

b) ¿Cuál es el número condicional esperado de pernos clasificados como bajos, dado que 16 pernos se clasifican como altos?

c) ¿El número de pernos clasificados como altos y como bajos son variables aleatorias independientes?

6-100. Para evaluar el soporte técnico de un fabricante de computadoras, se registra el número de timbrazos antes de que un representante de servicio conteste la llamada. Históricamente, 70% de las llamadas se contestan en dos timbrazos o menos, 25% se contestan en tres o cuatro timbrazos, y el resto de las llamadas requiere cinco timbrazos o más. Suponga que se hacen 10 llamadas a este fabricante y que las llamadas son independientes.

a) ¿Cuál es la probabilidad de que ocho llamadas se respondan en dos timbrazos o menos, una llamada se responda en tres o cuatro timbrazos, y una llamada requiera cinco timbrazos o más?

b) ¿Cuál es la probabilidad de que las 10 llamadas se contesten en cuatro timbrazos o menos?

c) ¿Cuál es el número esperado de llamadas que se contestan en cuatro timbrazos o menos?

6-101. Continuación del ejercicio 6-100.

a) ¿Cuál es la distribución condicional del número de llamadas que requieren cinco timbrazos o más, dado que ocho llamadas se contestan en dos timbrazos o menos?

b) ¿Cuál es el número condicional esperado de llamadas que requieren cinco timbrazos o más, dado que ocho llamadas se contestan en dos timbrazos o menos?

c) ¿El número de llamadas que se contestan en dos timbrazos o menos y el número de llamadas que requieren cinco timbrazos o más son variables aleatorias independientes?

6-102. Determine el valor de c tal que la función $f(x, y) = cx^2y$ para $0 < x < 3$ y $0 < y < 2$ satisfaga las propiedades de una función de densidad de probabilidad conjunta.

6-103. Continuación del ejercicio 6-102. Determine lo siguiente:

a) $P(X < 1, Y < 1)$

b) $P(X < 2.5)$

c) $P(1 < Y < 2.5)$

d) $P(X > 2, 1 < Y < 1.5)$

e) $E(X)$

f) $E(Y)$

6-104. Continuación del ejercicio 6-102.

a) Determine la distribución de probabilidad marginal de la variable aleatoria X.

b) Determine la distribución de probabilidad condicional de Y, dado que $X = 1$.

c) Determine la distribución de probabilidad condicional de X, dado que $Y = 1$.

6-105. La distribución conjunta de las variables aleatorias continuas X, Y y Z es constante en la región $x^2 + y^2 \leq 1$, $0 < z < 4$.

a) Determine $P(X^2 + Y^2 \leq 0.5)$

b) Determine $P(X^2 + Y^2 \leq 0.5, Z < 2)$

c) ¿Cuál es la función de densidad de probabilidad conjunta de X y Y, dado que $Z = 1$?

d) ¿Cuál es la función de densidad de probabilidad marginal de X?

6-106. Continuación del ejercicio 6-105.

a) Determine la media condicional de Z, dado que $X = 0$ y $Y = 0$.

b) En general, determine la media condicional de Z, dado que $X = x$ y $Y = y$.

6-107. Suponga que X y Y son variables aleatorias uniformes, continuas e independientes para $0 < x < 1$ y $0 < y < 1$. Use la función de densidad de probabilidad conjunta para determinar la probabilidad de que $|X - Y| < 0.5$.

6-108. El tiempo de vida útil de seis componentes principales de una copiadora es una variable aleatoria exponencial independiente con media de 8 000 horas, 10 000 horas, 10 000

horas, 20 000 horas, 20 000 horas y 25 000 horas, respectivamente.

a) ¿Cuál es la probabilidad de que el tiempo de vida útil de todos los componentes exceda 5 000 horas?

b) ¿Cuál es la covarianza entre el tiempo de vida útil de los componentes de 5 000 y 25 000 horas?

6-109. Los problemas de contaminación en la fabricación de semiconductores pueden dar como resultado un defecto funcional, un defecto menor o ningún defecto en el producto final. Suponga que 20%, 50% y 30% de los problemas de contaminación dan como resultado defectos funcionales, menores y ningún defecto, respectivamente. Suponga que los efectos de 10 problemas de contaminación son independientes.

a) ¿Cuál es la probabilidad de que los 10 problemas de contaminación den como resultado dos defectos funcionales y cinco defectos menores?

b) ¿Cuál es la distribución del número de problemas de contaminación que no dan como resultado un defecto?

c) ¿Cuál es el número esperado de problemas de contaminación que dan como resultado ningún defecto?

6-110. El peso de los ladrillos de adobe usados en la construcción tiene una distribución normal con una media de 3 libras y una desviación estándar de 0.25 libras. Suponga que el peso de los ladrillos es independiente y que se selecciona una muestra aleatoria de 25 ladrillos. ¿Cuál es la probabilidad de que el peso medio de la muestra sea menor que 2.95 libras?

6-111. El largo y ancho en pulgadas de los tableros usados en puertas interiores se denotan por X y Y, respectivamente. Suponga que X y Y son variables aleatorias uniformes, continuas e independientes para $17.75 < x < 18.25$ y $4.75 < y < 5.25$, respectivamente. Integrando la función de densidad de probabilidad conjunta en la región adecuada, determine la probabilidad de que el área de un tablero exceda 90 pulgadas cuadradas.

6-112. El peso de un dulce pequeño tiene una distribución normal con una media de 0.1 onzas y una desviación estándar de 0.01 onzas. Suponga que se colocan 16 dulces en un paquete y que los pesos son independientes.

a) ¿Cuál es la media y la varianza del peso neto del paquete?

b) ¿Cuál es la probabilidad de que el peso neto del paquete sea menor que 1.6 onzas?

c) Si se colocan 17 dulces en cada paquete, ¿cuál es la probabilidad de que el peso neto de un paquete sea menor que 1.6 onzas?

6-113. El tiempo para que el sistema automatizado de una bodega localice una pieza tiene una distribución normal con una media de 45 segundos y una desviación estándar de 30 segundos. Suponga que se hacen pedidos independientes de 10 piezas.

a) ¿Cuál es la probabilidad de que el tiempo promedio para localizar las 10 piezas exceda 60 segundos?

b) ¿Cuál es la probabilidad de que el tiempo total para localizar las 10 piezas exceda 600 segundos?

6-114. Un ensamblaje mecánico usando en el motor de un automóvil contiene cuatro componentes principales. Los pesos de los componentes son independientes y tienen una distribución normal, con las siguientes medias y desviaciones estándar (en onzas).

Componente	Media	Desviación estándar
cubierta izquierda	4	0.4
cubierta derecha	5.5	0.5
ensamblaje de rodamientos	10	0.2
ensamblaje de pernos	8	0.5

a) ¿Cuál es la probabilidad de que el peso de un ensamblaje exceda 29.5 onzas?

b) ¿Cuál es la probabilidad de que el peso medio de ocho ensamblajes independientes exceda 29 onzas?

6-115. Suponga que X y Y tienen una distribución normal de dos variables con $\sigma_X = 4$, $\sigma_Y = 1$, $\mu_X = 4$, $\mu_Y = 4$, y $\sigma = -0.2$. Trace una gráfica de contorno aproximada de la función de densidad de probabilidad conjunta.

6-116. Si

$$f_{XY}(x, y) = \frac{1}{1.2\pi}\exp\left\{\frac{-1}{0.72}\left[(x-1)^2 - 1.6(x-1)(y-2) + (y-2)^2\right]\right\}$$

determine $E(X)$, $E(Y)$, $V(X)$, $V(Y)$ y ρ reordenando los parámetros en la función de densidad de probabilidad conjunta.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

6-117. Demuestre la regla de Chebyshev. Defina la variable aleatoria Y de la siguiente manera:

$$Y = \begin{cases} 1 & \text{si } |X - \mu| \geq c\sigma \\ 0 & \text{en caso contrario} \end{cases}$$

a) Determine $E(Y)$

b) Muestre que $(X - \mu)^2 \geq (X - \mu)^2 Y \geq c^2\sigma^2 Y$

c) Utilizando el inciso b), demuestre que $E[(X - \mu)^2] \geq c^2\sigma^2 E[Y]$

d) Utilizando el inciso c), complete la deducción de la desigualdad de Chebyshev.

6-118. Demuestre que si $X_1, X_2, \ldots, X_p$ son variables aleatorias continuas e independientes, entonces $P(X_1 \in A_1, X_2 \in A_2, \ldots, X_p \in A_p) = P(X_1 \in A_1)P(X_2 \in A_2) \ldots P(X_p \in A_p)$ para cualesquiera regiones $A_1, A_2, \ldots, A_p$ en el rango de $X_1, X_2, \ldots, X_p$, respectivamente.

6-119. Demuestre que si $X_1, X_2, \ldots, X_p$ son variables aleatorias independientes y $Y = c_1X_1 + c_2X_2 + \ldots + c_pX_p$, entonces

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p)$$

Puede suponerse que las variables aleatorias son continuas.

6-120. Suponga que la función de densidad de probabilidad conjunta de las variables aleatorias continuas X y Y es constante en el rectángulo $0 < x < a$, $0 < y < b$. Demuestre que X y Y son independientes.

6-121. Suponga que el rango de las variables aleatorias continuas X y Y es $0 < x < a$ y $0 < y < b$. Suponga asimismo que la función de densidad de probabilidad conjunta $f_{XY}(x, y) = g(x)h(y)$, donde $g(x)$ es función solamente de x y $h(y)$ es función solamente de y. Muestre que X y Y son independientes.

7 Estimación de parámetros

PLAN GENERAL DEL CAPÍTULO

7-1 INFERENCIA ESTADÍSTICA

7-2 MUESTREO ALEATORIO

7-3 PROPIEDADES DE LOS ESTIMADORES

- 7-3.1 Estimadores insesgados
- 7-3.2 Varianza de un estimador puntual
- 7-3.3 Error estándar: reporte de una estimación puntual
- 7-3.4 Estimación *bootstrap* (cargador inicial) del error estándar
- 7-3.5 Error cuadrado medio de un estimador

7-4 MÉTODO DE MÁXIMA VEROSIMILITUD

7-5 DISTRIBUCIONES DE MUESTREO

7-6 DISTRIBUCIONES DE MUESTREO DE MEDIAS

7-7 INTRODUCCIÓN A LOS INTERVALOS DE CONFIANZA

7-1 INFERENCIA ESTADÍSTICA

El campo de la inferencia estadística se compone de los métodos que se utilizan para tomar decisiones o sacar conclusiones acerca de una **población**. Estos métodos emplean la información contenida en una **muestra** de la población para sacar conclusiones (véase la figura 7-1, donde se muestra la relación entre una población y una muestra). En este capítulo se inicia el estudio de los métodos estadísticos que se emplean para hacer inferencias y tomar decisiones.

La inferencia estadística puede dividirse en dos áreas principales: **estimación de parámetros** y **prueba de hipótesis**. Como ejemplo de un problema de estimación de parámetros, suponga que un ingeniero está analizando la resistencia a la tensión de un componente usado en un chasis automotriz. Puesto que la variabilidad de la resistencia a la tensión se encuentra presente

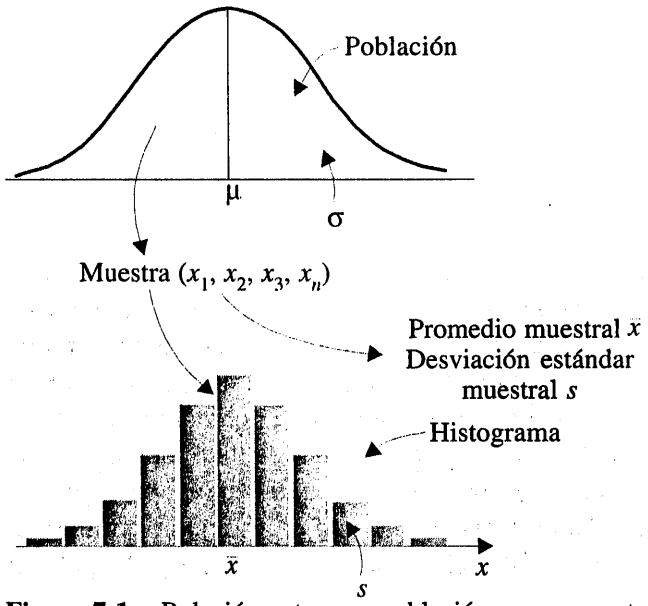


Figura 7-1 Relación entre una población y una muestra.

de manera natural entre los componentes individuales, debido a las diferencias en los lotes de materias primas, los procesos de manufactura y los procedimientos de medición (entre otros factores), al ingeniero le interesa hacer una estimación de la media de la resistencia a la tensión de los componentes. En la práctica, el ingeniero usará datos muestrales para calcular el número que, en cierto sentido, sea un valor (o una conjetura) razonable de la media real. A este número se le llama **estimación puntual**. Más adelante se verá que es posible determinar la precisión de la estimación.

Considérese ahora una situación en la que pueden usarse dos temperaturas de reacción diferentes en un proceso químico, digamos t_1 y t_2. El ingeniero conjetura que t_1 produce rendimientos más altos que t_2. La prueba de hipótesis estadística es el marco de referencia para resolver problemas de este tipo. En este caso, la hipótesis sería que el rendimiento medio utilizando la temperatura t_1 es mayor que el rendimiento medio utilizando la temperatura t_2. Obsérvese que no se hace énfasis en la estimación de los rendimientos; más bien, la atención se centra en sacar conclusiones acerca de una hipótesis propuesta.

El capítulo se inicia con el análisis de las **muestras aleatorias**, un concepto clave de la inferencia estadística, y después se presentan los métodos para la estimación de parámetros. Se exponen asimismo algunas de las propiedades más importantes de los estimadores y se indica cómo evaluar la precisión de la estimación de un parámetro con un **intervalo de confianza**. Por último, se concluye con una introducción a las **distribuciones de muestreo**, las cuales constituyen el fundamento de las técnicas para probar hipótesis y las técnicas para la estimación de parámetros que se presentan en los capítulos 8 y 9.

7-2 MUESTREO ALEATORIO

En la mayoría de los problemas de estadística es necesario usar una muestra de las observaciones de la población de interés para sacar conclusiones acerca de la población. Estos conceptos ya se comentaron; sin embargo, se presentan ahora las definiciones formales de algunos de estos términos.

Definición

Una **población** consiste en la totalidad de las observaciones que son motivo de interés.

En cualquier problema particular, la población puede ser pequeña, grande, pero finita, o infinita. Al número de observaciones que hay en la población se le llama el **tamaño** de la población. Por ejemplo, el número de botellas con un llenado incompleto que se producen en un día en una embotelladora de refrescos es una población de tamaño finito. Las observaciones obtenidas al medir el nivel diario de monóxido de carbono es una población de tamaño infinito. Es común usar una distribución de probabilidad como modelo de una población. Por ejemplo, un ingeniero de estructuras automotrices podría considerar que la población de resistencias a la tensión de un elemento estructural del chasis sigue una distribución normal con media μ y varianza σ^2. Podría hacerse referencia a ésta como una **población normal** o población distribuida normalmente.

En la mayoría de los problemas de inferencia estadística es imposible o poco práctico observar la población completa. Por ejemplo, quizá no se pruebe la resistencia a la tensión de todos los elementos estructurales del chasis porque se necesitaría una gran cantidad de tiempo y sería costoso. Además, algunos (quizá muchos) de estos elementos estructurales aún no existen en el momento de tomar una decisión, por lo que la población debe considerarse, en gran medida, **conceptual**. Por lo tanto, dependeremos de un subconjunto de observaciones de la población que nos ayude a tomar decisiones sobre la población.

Definición

Una **muestra** es un subconjunto de observaciones que se seleccionan de una población.

Para que las inferencias sean válidas, la muestra debe ser representativa de la población. Muchas veces resulta tentador seleccionar como muestra las observaciones que están más a la mano o ejercitar un juicio para seleccionar una muestra. Estos procedimientos suelen introducir *sesgos* en la muestra y, como resultado, el parámetro de interés se subestimará (o sobreestimará) consistentemente al utilizarla. Además, el comportamiento de una muestra dirigida no puede describirse estadísticamente. Para evitar estas dificultades, es deseable seleccionar una **muestra aleatoria** que resulte de algún mecanismo de azar. Por consiguiente, la selección de una muestra es un experimento aleatorio, y cada observación de la muestra es el valor observado de una variable aleatoria. Las observaciones que hay en la población determinan la distribución de probabilidad de la variable aleatoria.

Para definir una muestra aleatoria, sea X una variable aleatoria que representa el resultado de una selección de una observación de la población. Sea que $f(x)$ denote la función de densidad de probabilidad de X. Suponga que cada observación de la muestra se obtiene de manera independiente, bajo las mismas condiciones. Es decir, las observaciones para la muestra se obtienen

observando X de manera independiente y bajo condiciones sin cambios, digamos, n veces. Sea que X_i denote la variable aleatoria que representa la repetición i. Entonces $X_1, X_2, ..., X_n$ constituyen una muestra aleatoria y los valores numéricos obtenidos se denotan como $x_1, x_2, ..., x_n$. Las variables aleatorias de una muestra aleatoria son independientes y tienen la misma distribución de probabilidad $f(x)$ debido a las condiciones idénticas bajo las cuales se obtiene cada observación. Es decir, la función de densidad de probabilidad marginal de $X_1, X_2, ..., X_n$ es $f(x_1), f(x_2), ..., f(x_n)$, respectivamente, y por la independencia, la función de densidad de probabilidad conjunta de la muestra aleatoria es $f_{X_1X_2...X_n}(x_1, x_2, ..., x_n) = f(x_1)\,f(x_2)\,...\,f(x_n)$.

Definición

Las variables aleatorias $(X_1, X_2, ..., X_n)$ son una muestra aleatoria de tamaño n si a) las X_i son variables aleatorias independientes, y b) cada X_i tiene la misma distribución de probabilidad.

Para ilustrar esta definición, suponga que se está investigando la vida de servicio efectivo de un componente electrónico usado en un marcapasos cardiaco y que la vida del componente está distribuida normalmente. Se esperaría entonces que cada una de las observaciones de la vida del componente $X_1, X_2, ..., X_n$, en una muestra aleatoria de n componentes, sea una variable aleatoria independiente con exactamente la misma distribución normal. Después de recabar los datos, los valores numéricos de la vida útil observada se denotan como $x_1, x_2, ..., x_n$.

El objetivo principal al tomar una muestra aleatoria es obtener información acerca de los parámetros desconocidos de la población. Suponga, por ejemplo, que se quiere llegar a una conclusión acerca de la proporción de la población de Estados Unidos que prefiere una marca de refresco particular. Sea que p represente el valor desconocido de esta proporción. No es práctico preguntarle a cada individuo de la población para determinar el valor real de p. Para hacer una inferencia respecto a la proporción real p, un procedimiento más razonable sería seleccionar una muestra aleatoria (de un tamaño apropiado) y usar la proporción observada $\hat{p}$ de las personas de esta muestra que prefieren la marca de refresco.

La proporción muestral, $\hat{p}$, se calcula dividiendo el número de individuos de la muestra que prefieren la marca de refresco entre el tamaño total de la muestra n. Por tanto, $\hat{p}$ es una función de los valores observados en la muestra aleatoria. Puesto que es posible seleccionar muchas muestras de una población, el valor de $\hat{p}$ variará de una muestra a otra. Es decir, $\hat{p}$ es una variable aleatoria. A esta variable aleatoria se le llama **estadístico**.

Definición

Un **estadístico** es cualquier función de las observaciones de una muestra aleatoria.

Nos hemos encontrado ya con estadísticos. Por ejemplo, si $X_1, X_2, ..., X_n$ es una muestra aleatoria de tamaño n, entonces la media muestral $\overline{X}$, la varianza muestral S^2, y la desviación

estándar muestral S son estadísticos. Estos estadísticos se usan ampliamente en el proceso de sacar conclusiones acerca de poblaciones con base en datos muestrales.

Puesto que un estadístico es una variable aleatoria, tiene una distribución de probabilidad. A la distribución de probabilidad de un estadístico se le llama **distribución de muestreo**. El concepto de distribución de muestreo es muy importante y se revisará y ejemplificará más adelante en este capítulo.

Una aplicación muy importante de los estadísticos es para obtener **estimaciones puntuales** de parámetros tales como la media poblacional y la varianza poblacional. Cuando se analizan problemas de inferencia, es conveniente tener un símbolo general que represente el parámetro de interés. Se usará la letra griega θ (theta) para representar el parámetro. El objetivo de una estimación puntual es seleccionar, con base en los datos muestrales, un solo número que sea el valor más recomendable de θ. Se usará el valor numérico de un estadístico muestral como la estimación puntual.

En general, si X es una variable aleatoria con distribución de probabilidad $f(x)$, caracterizada por el parámetro desconocido θ, y si $X_1, X_2, ..., X_n$ es una muestra aleatoria de tamaño n de X, entonces al estadístico $\hat{\Theta} = h(X_1, X_2, ..., X_n)$ se le llama un **estimador puntual** de θ. Obsérvese que $\hat{\Theta}$ es una variable aleatoria, ya que es una función de variables aleatorias. Después de seleccionar la muestra, $\hat{\Theta}$ toma un valor numérico particular $\hat{\theta}$ llamado la **estimación puntual** de θ.

Definición

Una **estimación puntual** de un parámetro poblacional θ es el valor numérico particular $\hat{\theta}$ de un estadístico $\hat{\Theta}$.

Como ejemplo, suponga que la variable aleatoria X tiene una distribución normal con una media desconocida μ. La media muestral es un estimador puntual de la media poblacional desconocida μ. Es decir, $\hat{\mu} = \overline{X}$. Después de seleccionar la muestra, el valor numérico $\bar{x}$ es la estimación puntual de μ. Por tanto, si $x_1 = 25$, $x_2 = 30$, $x_3 = 29$, y $x_4 = 31$, entonces la estimación puntual de μ es

$$\bar{x} = \frac{25 + 30 + 29 + 31}{4} = 28.75$$

De manera similar, si la varianza poblacional σ^2 también es desconocida, un estimador puntual de σ^2 es la varianza muestral S^2, y al valor numérico $s^2 = 6.9$, calculado a partir de los datos muestrales, se le llama la estimación puntual de σ^2.

En la ingeniería son comunes los problemas de estimación. Con frecuencia es necesario estimar:

- La media μ de una sola población.
- La varianza σ^2 (o la desviación estándar σ) de una sola población.
- La proporción p de elementos de una población que pertenecen a una clase de interés.
- La diferencia en las medias de dos poblaciones, $\mu_1 - \mu_2$.
- La diferencia en las proporciones de dos poblaciones, $p_1 - p_2$.

Estimaciones razonables de estos parámetros son las siguientes:

- Para μ, la estimación es $\hat{\mu} = \bar{x}$, la media muestral.
- Para σ^2, la estimación es $\hat{\sigma}^2 = s^2$, la varianza muestral.
- Para p, la estimación es $\hat{p} = x/n$, la proporción muestral, donde x es el número de elementos de una muestra aleatoria de tamaño n que pertenecen a la clase de interés.
- Para $\mu_1 - \mu_2$, la estimación es $\hat{\mu}_1 - \hat{\mu}_2 = \bar{x}_1 - \bar{x}_2$, la diferencia entre las medias muestrales de dos muestras aleatorias independientes.
- Para $p_1 - p_2$, la estimación es $\hat{p}_1 - \hat{p}_2$, la diferencia entre dos proporciones muestrales calculadas a partir de dos muestras aleatorias independientes.

Puede haber varias opciones diferentes para el estimador puntual de un parámetro. Por ejemplo, si quiere estimarse la media de una población, podrían considerarse como estimadores puntuales la media muestral, la mediana muestral o quizá el promedio de las observaciones menor y mayor de la muestra. Para decidir cuál de los estimadores puntuales de un parámetro particular es el más adecuado, es necesario examinar sus propiedades estadísticas y desarrollar algunos criterios para comparar estimadores.

7-3 PROPIEDADES DE LOS ESTIMADORES

7-3.1 Estimadores insesgados

Un estimador deberá estar en algún sentido "cerca" del valor real del parámetro desconocido. En términos formales, se dice que $\hat{\Theta}$ es un estimador insesgado de θ si el valor esperado de Θ es igual a θ. Esto es equivalente a decir que la media de la distribución de probabilidad de Θ (o la media de la distribución de muestreo de $\hat{\Theta}$) es igual a θ.

Definición

El estimador puntual $\hat{\Theta}$ es un **estimador insesgado** del parámetro θ si

$$E(\hat{\Theta}) = \theta \tag{7-1}$$

Si el estimador no es insesgado, entonces, a la diferencia

$$E(\hat{\Theta}) - \theta \tag{7-2}$$

se le llama el **sesgo** del estimador $\hat{\Theta}$.

Cuando un estimador es insesgado, el sesgo es cero; es decir, $E(\hat{\Theta}) - \theta = 0$.

EJEMPLO 7-1

Suponga que X es una variable aleatoria con media μ y varianza σ^2. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de tamaño n de la población representada por X. Demuestre que la media muestral $\bar{X}$ y la varianza muestral S^2 son estimadores insesgados de μ y σ^2, respectivamente.

Considérese primero la media muestral. En la ecuación 6-38 del capítulo 6 se estableció que $E(\bar{X}) = \mu$. Por lo tanto, la media muestral $\bar{X}$ es un estimador insesgado de la media poblacional μ.

Considérese ahora la varianza muestral. Se tiene

$$E(S^2) = E\left(\frac{\sum_{i=1}^{n} = (X_i - \bar{X})^2}{n-1}\right) = \frac{1}{n-1} E \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$= \frac{1}{n-1} E \sum_{i=1}^{n} (X_i^2 + \bar{X}^2 - 2\bar{X}X_i) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)\right)$$

La última igualdad se sigue de la ecuación 6-36 del capítulo 6. Sin embargo, puesto que $E(X_i^2) = \mu^2 + \sigma^2$ y $E(\bar{X}^2) = \mu^2 + \sigma^2/n$, se tiene

$$E(S^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n} (\mu^2 + \sigma^2) - n(\mu^2 + \sigma^2/n)\right]$$

$$= \frac{1}{n-1}(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2)$$

$$= \sigma^2$$

Por lo tanto, la varianza muestral S^2 es un estimador insesgado de la varianza poblacional σ^2. Sin embargo, puede demostrarse que la desviación estándar muestral S es un estimador sesgado de la desviación estándar poblacional. Para muestras grandes este sesgo es despreciable.

En ocasiones hay varios estimadores insesgados del parámetro poblacional. Por ejemplo, suponga que se toma una muestra aleatoria de tamaño $n = 10$ de una población normal y se obtienen los datos $x_1 = 12.8$, $x_2 = 9.4$, $x_3 = 8.7$, $x_4 = 11.6$, $x_5 = 13.1$, $x_6 = 9.8$, $x_7 = 14.1$, $x_8 = 8.5$, $x_9 = 12.1$, $x_{10} = 10.3$. Entonces la media muestral es

$$\bar{x} = \frac{12.8 + 9.4 + 8.7 + 11.6 + 13.1 + 9.8 + 14.1 + 8.5 + 12.1 + 10.3}{10}$$

$$= 11.04$$

la mediana muestral es

$$\bar{x} = \frac{10.3 + 11.6}{2} = 10.95$$

y una media recortada al 10% (la cual se obtiene descartando el 10% más pequeño y más grande de la muestra antes de calcular el promedio) es

$$\bar{x}_{rec(10)} = \frac{8.7 + 9.4 + 9.8 + 10.3 + 11.6 + 12.1 + 12.8 + 13.1}{8}$$

Puede demostrarse que todas éstas son estimaciones insesgadas de μ. Puesto que no existe un estimador insesgado único, no es posible confiar tan sólo en la propiedad de la ausencia de sesgo para seleccionar el estimador. Se necesita un método para escoger entre los estimadores insesgados. En la sección siguiente se sugiere un método.

7-3.2 Varianza de un estimador puntual

Suponga que $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son estimadores insesgados de θ. Esto indica que la distribución de cada uno de los estimadores tiene su centro en el valor real de θ. Sin embargo, las varianzas de estas distribuciones pueden ser diferentes. En la figura 7-2 se ilustra esta situación. Puesto que $\hat{\Theta}_1$ tiene una varianza menor que $\hat{\Theta}_2$, el estimador $\hat{\Theta}_1$ tiene mayores posibilidades de producir una estimación que esté próxima al valor real de θ. Un principio lógico de la estimación, cuando se escoge entre varios estimadores, es elegir el estimador que tenga la varianza mínima.

Definición

Si se consideran todos los estimadores insesgados de θ, al que tiene la varianza menor se le llama el **estimador insesgado de varianza mínima** (MVUE).

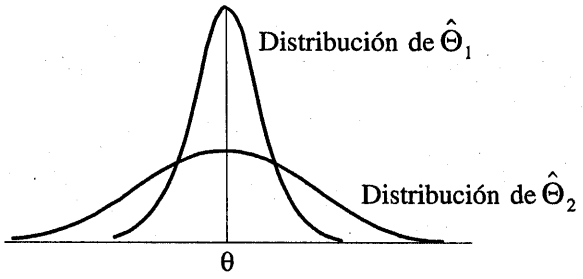


Figura 7-2 Las distribuciones de muestreo de dos estimadores insesgados $\hat{\Theta}_1$ y $\hat{\Theta}_2$.

En cierto sentido, el MVUE es el estimador que tiene las mayores posibilidades de producir una estimación $\hat{\theta}$ que esté próxima al valor real de θ. Se ha podido desarrollar una metodología para identificar el MVUE en muchas situaciones prácticas. Si bien esta metodología está fuera del alcance del presente libro, se presenta un resultado muy importante relacionado con la distribución normal.

Teorema 7-1

Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de tamaño n de una distribución normal con media μ y varianza σ^2, entonces la media muestral $\overline{X}$ es el MVUE de μ.

Incluso en situaciones en las que no se sabe si existe un MVUE podría usarse el principio de la varianza mínima para elegir entre los posibles estimadores. Suponga, por ejemplo, que se quiere estimar la media de una población (no necesariamente una población *normal*). Se tiene una muestra aleatoria de n observaciones $X_1, X_2, \ldots, X_n$ y se quieren comparar dos posibles estimadores de μ: la media muestral $\overline{X}$ y una observación particular de la muestra, digamos, X_i. Obsérvese que tanto $\overline{X}$ como X_i son estimadores insesgados de μ; para la media muestral se tiene $V(\overline{X}) = \sigma^2/n$, por la ecuación 6-38, y la varianza de cualquier observación es $V(X_i) = \sigma^2$. Puesto que $V(\overline{X}) < V(X_i)$ para cualquier tamaño de la muestra $n \geq 2$, se concluiría que la media muestral es un mejor estimador de μ que una observación simple X_i.

7-3.3 Error estándar: reporte de una estimación puntual

Cuando se reporta el valor numérico o estimación puntual de un parámetro, suele ser conveniente dar cierta idea de la precisión de la estimación. La medida de la precisión que se usa generalmente es el error estándar del estimador que se ha usado.

Definición

El **error estándar** de un estimador $\hat{\Theta}$ es su desviación estándar, dada por $\sigma_{\hat{\Theta}} = \sqrt{V(\hat{\Theta})}$. Si el error estándar incluye parámetros desconocidos que pueden estimarse, entonces la sustitución de dichos valores en $\sigma_{\hat{\Theta}}$ produce un **error estándar estimado**, denotado por $\hat{\sigma}_{\hat{\Theta}}$.

En ocasiones el error estándar estimado se denota por $s_{\hat{\Theta}}$ o $se(\hat{\Theta})$.

Suponga que se hace un muestreo de una distribución normal con media μ y varianza σ^2. Entonces la distribución de $\overline{X}$ es normal con media μ y varianza σ^2/n, por lo que el error estándar de $\overline{X}$ es

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

Si no se conoce σ, pero se sustituye la desviación estándar muestral S en la ecuación anterior, entonces el error estándar estimado de $\bar{X}$ es

$$\hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}}$$

Cuando el estimador sigue una distribución normal, como en la situación anterior, puede tenerse una seguridad razonable de que el valor real del parámetro está entre dos errores estándar de la estimación. Puesto que muchos estimadores puntuales tienen una distribución normal (al menos aproximadamente) para n grande, este resultado es muy útil. Incluso en casos en que el estimador puntual no tiene una distribución normal puede afirmarse que, en tanto el estimador sea insesgado, la estimación del parámetro tendrá una desviación del valor real de hasta cuatro errores estándar a lo sumo 6% de las veces. Así, una afirmación muy conservadora es que el valor real del parámetro difiere de la estimación puntual en a lo sumo cuatro errores estándar.

EJEMPLO 7-2

Un artículo del *Journal of Heat Transfer* (Trans. ASME, Secc. C, 96, p. 59) describía un nuevo método para medir la conductividad térmica del hierro Armco. Utilizando una temperatura de 100°F y una alimentación de energía de 550 W, se obtuvieron las siguientes 10 mediciones de la conductividad térmica (en Btu/hr-pie-°F):

41.60, 41.48, 42.34, 41.95, 41.86,
42.18, 41.72, 42.26, 41.81, 42.04

Una estimación puntual de la conductividad térmica media a 100°F y 550 W es la media muestral o

$$\bar{x} = 41.924 \text{ Btu/hr-pie-°F}$$

El error estándar de la media muestral es $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, y puesto que no se conoce σ, puede sustituirse con la desviación estándar muestral $s = 0.284$ para obtener el error estándar estimado de $\bar{X}$ como

$$\hat{\sigma}_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{0.284}{\sqrt{10}} = 0.0898$$

Obsérvese que el error estándar es aproximadamente 0.2% de la media muestral, lo cual implica que se ha obtenido una estimación puntual relativamente precisa de la conductividad térmica. Si puede suponerse que la conductividad térmica tiene una distribución normal, entonces dos veces el error estándar es $2\hat{\sigma}_{\bar{X}} = 2(0.0898) = 0.1796$, y se tiene una gran confianza en que la conductividad térmica media real está en el intervalo 41.924 ± 0.1756, o entre 41.744 y 42.104.

7-3.4 Estimación *bootstrap* (cargador inicial) del error estándar

Hay situaciones en las que no se conoce el error estándar del estimador puntual. Por lo general, se trata de casos en que la forma de $\hat{\Theta}$ es complicada, y es difícil aplicar los operadores convencionales de expectativa y varianza. Una técnica de computadora llamada ***bootstrap*** (**cargador inicial**) que se desarrolló en años recientes puede usarse para este problema.

Suponga que se está haciendo el muestreo de una población que puede modelarse con la distribución de probabilidad $f(x; \theta)$. La muestra aleatoria da como resultado los valores numéricos $x_1, x_2, ..., x_n$ y se obtiene $\hat{\theta}$ como la estimación puntual de θ. Ahora se usaría una computadora para obtener *muestras bootstrap* de la distribución $f(x; \hat{\theta})$, y para cada una de estas muestras se calcula la estimación *bootstrap* $\hat{\theta}^*$ de θ. Se obtiene así:

Muestra *bootstrap*	Observaciones	Estimación *bootstrap*
1	$x_1^*, x_2^*, ..., x_n^*$	$\hat{\theta}_1^*$
2	$x_1^*, x_2^*, ..., x_n^*$	$\hat{\theta}_2^*$
.	.	.
.	.	.
.	.	.
B	$x_1^*, x_2^*, ..., x_n^*$	$\hat{\theta}_B^*$

Por lo general se toman $B = 100$ o 200 de estas muestras *bootstrap*. Sea $\bar{\theta}^* = (1/B) \sum_{i=1}^{B} \hat{\theta}_i^*$ la media muestral de las estimaciones *bootstrap*. La estimación *bootstrap* del error estándar de $\hat{\Theta}$ es simplemente la desviación estándar de las $\hat{\theta}_i^*$, o

$$s_{\hat{\vartheta}} = \sqrt{\frac{\sum_{i=1}^{B} (\hat{\theta}_i^* - \bar{\theta}^*)^2}{B-1}} \tag{7-3}$$

En la literatura *bootstrap*, es común reemplazar el término $B - 1$ de la ecuación 7-3 con B. Sin embargo, para los valores grandes que suelen emplearse para B, hay poca diferencia en la estimación obtenida de $s_{\hat{\vartheta}}$.

EJEMPLO 7-3

El tiempo de falla de un módulo electrónico usado en el controlador del motor de un automóvil se prueba a alta temperatura a fin de acelerar el mecanismo de falla. El tiempo de falla tiene una distribución exponencial con parámetro λ desconocido. Se seleccionan ocho unidades al azar y se someten a prueba, con los tiempos de falla resultantes (en horas): $x_1 = 11.96$, $x_2 = 5.03$, $x_3 = 67.40$, $x_4 = 16.07$, $x_5 = 31.50$, $x_6 = 7.73$, $x_7 = 11.10$ y $x_8 = 22.38$. Ahora bien, la media de una distribución exponencial es $\mu = 1/\lambda$, por lo que $E(X) = 1/\lambda$, y el valor esperado del promedio muestral es $E(\bar{X}) = 1/\lambda$. Por lo tanto, una forma razonable de estimar λ es con $\hat{\lambda} = 1/\bar{X}$. Para la

muestra en cuestión, $\bar{x} = 21.65$, por lo que la estimación de λ es $\hat{\lambda} = 1/21.65 = 0.0462$. Para encontrar el error estándar *bootstrap* se obtendrían $B = 200$ muestras (por ejemplo) de $n = 8$ observaciones cada una, de una distribución exponencial con parámetro $\lambda = 0.0462$. Se presentan a continuación algunos de estos resultados:

Muestra *bootstrap*	Observaciones	Estimación *bootstrap*
1	8.01, 28.85, 14.14, 59.12, 3.11, 32.19, 5.26, 14.17	$\hat{\lambda}_1^* = 0.0485$
2	33.27, 2.10, 40.17, 32.43, 6.94, 30.66, 18.99, 5.61	$\hat{\lambda}_2^* = 0.0470$
.	.	.
.	.	.
.	.	.
200	40.26, 39.26, 19.59, 43.53, 9.55, 7.07, 6.03, 8.94	$\hat{\lambda}_{100}^* = 0.0459$

El promedio muestral de las $\hat{\lambda}_i^*$ (las estimaciones *bootstrap*) es 0.0513, y la desviación estándar de estas estimaciones *bootstrap* es 0.020. Por lo tanto, el error estándar *bootstrap* de $\hat{\lambda}$ es 0.020. Entonces, en este caso resulta que, al estimar el parámetro λ en una distribución exponencial, la varianza del estimador que se usó, $\hat{\lambda}$, es conocida. Cuando n es grande, $V(\hat{\lambda}) = \lambda^2/n$. Por lo tanto, el error estándar estimado de $\hat{\lambda}$ es $\sqrt{\hat{\lambda}^2/n} = \sqrt{(0.0462)^2/8} = 0.016$. Obsérvese que este resultado concuerda razonablemente con el error estándar *bootstrap*.

En ocasiones quiere aplicarse el *bootstrap* en situaciones en que no se conoce la forma de la distribución de probabilidad. En estos casos, se toman como *población* las n observaciones de la muestra y se seleccionan B muestras aleatorias de tamaño n cada una, con reemplazo, de esta población. Entonces puede aplicarse la ecuación 7-3 como se describió arriba. El libro de Efron y Tibshirani es una excelente introducción al *bootstrap*.

7-3.5 Error cuadrado medio de un estimador

En ocasiones es necesario usar un estimador sesgado. En tales casos, el error cuadrado medio del estimador puede ser importante. El **error cuadrado medio** de un estimador $\hat{\Theta}$ es el valor esperado de la diferencia al cuadrado entre $\hat{\Theta}$ y θ.

Definición

El **error cuadrado medio** de un estimador $\hat{\Theta}$ del parámetro θ se define como

$$\text{MSE}(\hat{\Theta}) = E(\hat{\Theta} - \theta)^2 \tag{7-4}$$

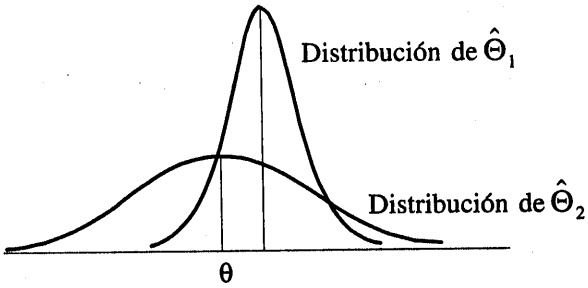


Figura 7-3 Un estimador sesgado $\hat{\Theta}_1$ que tiene una varianza menor que un estimador insesgado $\hat{\Theta}_2$.

El error cuadrado medio puede reescribirse de la siguiente manera:

$$\begin{aligned} \text{MSE}(\hat{\Theta}) &= E[\hat{\Theta} - E(\hat{\Theta})]^2 + [\theta - E(\hat{\Theta})]^2 \\ &= V(\hat{\Theta}) + (\text{bias})^2 \end{aligned}$$

Es decir, el error cuadrado medio de $\hat{\Theta}$ es igual a la varianza del estimador más el cuadrado del sesgo. Si $\hat{\Theta}$ es un estimador insesgado de θ, el error cuadrado medio de $\hat{\Theta}$ es igual a la varianza de $\hat{\Theta}$.

El error cuadrado medio es un criterio importante para comparar dos estimadores. Sean $\hat{\Theta}_1$ y $\hat{\Theta}_2$ dos estimadores del parámetro θ, y sean MSE($\hat{\Theta}_1$) y MSE($\hat{\Theta}_2$) los errores cuadrados medios de $\hat{\Theta}_1$ y $\hat{\Theta}_2$. Entonces, la **eficiencia relativa** de $\hat{\Theta}_2$ con respecto a $\hat{\Theta}_1$ se define como

$$\frac{\text{MSE}(\hat{\Theta}_1)}{\text{MSE}(\hat{\Theta}_2)} \tag{7-5}$$

Si esta eficiencia relativa es menor que uno, se concluiría que $\hat{\Theta}_1$ es un estimador más eficiente de θ que $\hat{\Theta}_2$, en el sentido de que tiene un error cuadrado medio menor.

En ocasiones se encuentra que los estimadores sesgados son preferibles a los insesgados debido a que tienen un error cuadrado medio menor. Es decir, puede conseguirse una reducción considerable en la varianza del estimador introduciendo una cantidad relativamente pequeña de sesgo. Siempre que la reducción en la varianza sea mayor que el cuadrado del sesgo, el resultado será un estimador mejorado desde el punto de vista del error cuadrado medio. Por ejemplo, en la figura 7-3 se muestra la distribución de probabilidad de un estimador sesgado $\hat{\Theta}_1$ que tiene una varianza menor que el estimador insesgado $\hat{\Theta}_2$. Una estimación basada en $\hat{\Theta}_1$ tendría mayores posibilidades de estar cerca del valor real de θ que una estimación basada en $\hat{\Theta}_2$. El análisis de regresión lineal (capítulos 10 y 11) es un campo donde se usan ocasionalmente estimadores sesgados.

Al estimador $\hat{\Theta}$ que tiene un error cuadrado medio que es menor o igual que el error cuadrado medio de cualquier otro estimador, para todos los valores del parámetro θ, se le llama estimador **óptimo** de θ. Los estimadores óptimos rara vez existen.

EJERCICIOS DE LA SECCIÓN 7-3

7-1. Suponga que se tiene una muestra aleatoria de tamaño $2n$ de una población denotada por X, y $E(X) = \mu$, $V(X) = \sigma^2$. Sean

$$\bar{X}_1 = \frac{1}{2n}\sum_{i=1}^{2n} X_i \quad \text{y} \quad \bar{X}_2 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

dos estimadores de μ. ¿Cuál es el mejor estimador de μ? Explique su elección.

7-2. Sea que $X_1, X_2, \ldots, X_7$ denote una muestra aleatoria de una población que tiene media μ y varianza σ^2. Considérense los siguientes estimadores de μ:

$$\hat{\Theta}_1 = \frac{X_1 + X_2 + \cdots + X_7}{7}$$

$$\hat{\Theta}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

a) ¿Alguno de los dos estimadores es insesgado?

b) ¿Cuál estimador es el mejor? ¿En qué sentido es mejor?

7-3. Suponga que $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son estimadores insesgados del parámetro θ. Se sabe que $V(\hat{\Theta}_1) = 10$ y $V(\hat{\Theta}_2) = 4$. ¿Cuál estimador es mejor y en qué sentido lo es?

7-4. Calcule la eficiencia relativa de los dos estimadores del ejercicio 7-2.

7-5. Calcule la eficiencia relativa de los dos estimadores del ejercicio 7-3.

7-6. Suponga que $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son estimadores del parámetro θ. Se sabe que $E(\hat{\Theta}_1) = \theta$, $E(\hat{\Theta}_2) = \theta/2$, $V(\hat{\Theta}_1) = 10$ y $V(\hat{\Theta}_2) = 4$. ¿Cuál estimador es mejor? ¿En qué sentido es mejor?

7-7. Suponga que $\hat{\Theta}_1$, $\hat{\Theta}_2$ y $\hat{\Theta}_3$ son estimadores del parámetro θ. Se sabe que $E(\hat{\Theta}_1) = E(\hat{\Theta}_2) = \theta$, $E(\hat{\Theta}_3) \neq \theta$, $V(\hat{\Theta}_1) = 12$, $V(\hat{\Theta}_2) = 10$ y $E(\hat{\Theta}_3 - \theta)^2 = 6$. Compare estos tres estimadores. ¿Por cuál se inclina el lector? ¿Por qué?

7-8. Sea que se tomen tres muestras aleatorias de tamaño $n_1 = 20$, $n_2 = 10$ y $n_3 = 8$ de una población con media μ y varianza σ^2. Sean S_1^2, S_2^2 y S_3^2 las varianzas muestrales. Demuestre que $S^2 = (20\,S_1^2 + 10\,S_2^2 + 8\,S_3^2)/38$ es un estimador insesgado de σ^2.

7-9. a) Demuestre que $\sum_{1=i}^{n}(X_i - \bar{X})^2/n$ es un estimador sesgado de σ^2.

b) Encuentre la cantidad de sesgo en el estimador.

c) ¿Qué ocurre con el sesgo cuando el tamaño de la muestra n se incrementa?

7-10. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n.

a) Demuestre que $\bar{X}^2$ es un estimador sesgado de μ^2.

b) Encuentre la cantidad de sesgo en este estimador.

c) ¿Qué ocurre con el sesgo cuando el tamaño de la muestra n se incrementa?

7-11. Se piensa que los defectos de un tablero metálico usado en la manufactura de automóviles sigue una distribución de Poisson. Se cuentan los defectos en 10 tableros con los siguientes resultados: $x_1 = 2, x_2 = 7, x_3 = 15, x_4 = 8, x_5 = 7$, $x_6 = 6$, $x_7 = 3$, $x_8 = 7$, $x_9 = 3$, $x_{10} = 4$. Encuentre una estimación puntual del parámetro de Poisson λ. Encuentre el error estándar *bootstrap* de λ.

7-12. Se extraen 10 observaciones al azar de una población con distribución de probabilidad desconocida. Las observaciones son $x_1 = 48$, $x_2 = 57$, $x_3 = 53$, $x_4 = 53$, $x_5 = 58$, $x_6 = 39$, $x_7 = 46$, $x_8 = 45$, $x_9 = 61$ y $x_{10} = 49$. Encuentre una media recortada al 10% de estos datos (remítase al ejercicio 2-38). Suponga que ésta es una estimación razonable de la media poblacional μ. Encuentre el error estándar *bootstrap* de la media recortada.

7-13. Considere de nuevo las observaciones del ejercicio 7-12. Calcule las desviaciones estándar muestrales de estas observaciones. Esta es una estimación razonable de la desviación estándarpoblacional σ. Encuentre el error estándar *bootstrap* de s.

7-14. Considere de nuevo las observaciones del ejercicio 7-12. Calcule el rango muestral de estas observaciones. Encuentre el error estándar *bootstrap* del rango muestral.

7-4 MÉTODO DE MÁXIMA VEROSIMILITUD

Uno de los mejores métodos para obtener un estimador puntual de un parámetro es el método de máxima verosimilitud. Como su nombre lo indica, el estimador será el valor del parámetro que maximiza la **función de verosimilitud**.

Definición

Suponga que X es una variable aleatoria con distribución de probabilidad $f(x; \theta)$, donde θ es un solo parámetro desconocido. Sean $x_1, x_2, \ldots, x_n$ los valores observados en una muestra aleatoria de tamaño n. Entonces la **función de verosimilitud** de la muestra es

$$L(\theta) = f(x_1; \theta) \cdot f(x_2; \theta) \cdot \cdots \cdot f(x_n; \theta) \tag{7-6}$$

Obsérvese que la función de verosimilitud es ahora una función exclusiva del parámetro desconocido θ. El **estimador de máxima verosimilitud** de θ es el valor de θ que maximiza la función de verosimilitud $L(\theta)$.

En el caso de una variable aleatoria discreta, la interpretación de la función de verosimilitud es clara. La función de verosimilitud de la muestra $L(\theta)$ es simplemente la probabilidad

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

Es decir, $L(\theta)$ es simplemente la probabilidad de obtener los valores muestrales $x_1, x_2, \ldots, x_n$. Por lo tanto, en el caso discreto, el estimador de máxima verosimilitud es un estimador que maximiza la probabilidad de la ocurrencia de los valores muestrales.

EJEMPLO 7-4

Sea X una variable aleatoria de Bernoulli. La función de masa de probabilidad es

$$f(x; p) = \begin{cases} p^x (1-p)^{1-x} & x = 0, 1 \\ 0 & \text{en caso contrario} \end{cases}$$

donde p es el parámetro que va a estimarse. La función de verosimilitud de una muestra de tamaño n es

$$L(p) = p^{x_1}(1-p)^{1-x_1} p^{x_2}(1-p)^{1-x_2} \cdots p^{x_n}(1-p)^{1-x_n}$$

$$= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$= p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i}$$

Se observa que si $\hat{p}$ maximiza a $L(p)$, entonces $\hat{p}$ también maximiza a $ln\ L(p)$. Por lo tanto,

$$ln\ L(p) = \left(\sum_{i=1}^{n} x_i\right) ln\ p + \left(n - \sum_{i=1}^{n} x_i\right) ln(1-p)$$

Entonces

$$\frac{d\ ln\ L(p)}{dp} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{\left(n - \sum_{i=1}^{n} x_i\right)}{1-p}$$

Al igualar con cero esta expresión y despejar p, se obtiene $\hat{p} = (1/n)\sum_{i=1}^{n} x_i$. Por lo tanto, el estimador de máxima verosimilitud de p es

$$\hat{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Suponga que este estimador se aplicó a la siguiente situación: se seleccionan al azar n unidades de una línea de producción, y cada unidad se clasifica como defectuosa (en cuyo caso se hace $x_i = 1$) o bien como no defectuosa (en cuyo caso se hace $x_i = 0$). Entonces $\sum_{i=1}^{n} x_i$ es el número de unidades defectuosas de la muestra, y $\hat{p}$ es la **proporción defectuosa de la muestra**. El parámetro p es la **proporción defectuosa de la población**; intuitivamente, parecería bastante razonable usar $\hat{p}$ como un estimador de p.

Aun cuando la interpretación de la función de verosimilitud dada arriba se limita al caso de una variable aleatoria discreta, el método de máxima verosimilitud puede extenderse fácilmente a una distribución continua. Se presentan a continuación dos ejemplos de la estimación de máxima verosimilitud para distribuciones continuas.

EJEMPLO 7-5

Sea que X tenga una distribución normal con μ desconocida y varianza σ^2 conocida. La función de verosimilitud de una muestra aleatoria de tamaño n, digamos $X_1, X_2, ..., X_n$, es

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/(2\sigma^2)}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-(1/2\sigma^2)\sum_{i=1}^{n}(X_i-\mu)^2}$$

Ahora bien,

$$ln\ L(\mu) = -(n/2)\, ln\,(2\pi\sigma^2) - (2\sigma^2)^{-1}\sum_{i=1}^{n}(x_i-\mu)^2$$

y

$$\frac{d\,ln\,L(\mu)}{d\mu} = (\sigma^2)^{-1}\sum_{i=1}^{n}(x_i - \mu)$$

Al igualar con cero esta última expresión y despejar μ, se obtiene

$$\hat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}$$

Por tanto, la media muestral es el estimador de máxima verosimilitud de μ.

EJEMPLO 7-6

Sea que X tenga una distribución exponencial con parámetro λ. La función de verosimilitud de una muestra aleatoria de tamaño n, digamos $X_1, X_2, ..., X_n$, es

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$= \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

La verosimilitud logarítmica es

$$ln\,L(\lambda) = n\,ln\,\lambda - \lambda\sum_{i=1}^{n} x_i$$

Entonces,

$$\frac{d\,ln\,L(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

y al igualar con cero este último resultado se obtiene

$$\lambda = n / \sum_{i=1}^{n} X_i = 1/\bar{X}$$

Por tanto, el estimador de máxima verosimilitud de λ es el recíproco de la media muestral.

Es sencillo ilustrar gráficamente cómo funciona el método de máxima verosimilitud. En la figura 7-4 se grafica el logaritmo de la función de verosimilitud para el parámetro exponencial del ejemplo 7-6, utilizando los $n = 8$ datos del ejemplo 7-3. En el ejemplo 7-3 se encontró que la estimación de λ fue $\hat{\lambda} = 0.0462$. Por el ejemplo 7-6 se sabe que ésta es una estimación de máxi-

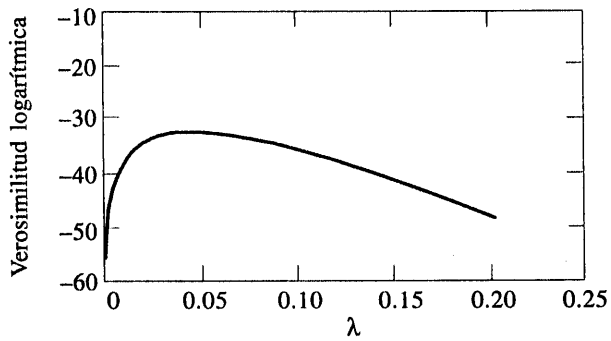


Figura 7-4 Verosimilitud logarítmica de la distribución exponencial, utilizando los datos del ejemplo 7-2.

ma verosimilitud. En la figura 7-4 se observa claramente que la función de verosimilitud logarítmica se maximiza con un valor de λ que es aproximadamente igual a 0.0462. Obsérvese que la función de verosimilitud logarítmica es relativamente plana en la región del máximo. Esto implica que la estimación del parámetro no es muy precisa. Si la estimación del parámetro fuera precisa, la función de verosimilitud logarítmica sería muy puntiaguda en el valor máximo. En este caso el tamaño de la muestra es relativamente pequeño, lo cual probablemente haya llevado a la imprecisión de la estimación.

El método de máxima verosimilitud puede usarse en situaciones en las que hay varios parámetros desconocidos, digamos $\theta_1, \theta_2, ..., \theta_k$, por estimar. En tales casos, la función de verosimilitud es una función de los k parámetros desconocidos $\theta_1, \theta_2, ..., \theta_k$, y los estimadores de máxima verosimilitud $\{\hat{\Theta}_i\}$ se encontrarían igualando con cero las k derivadas parciales $\partial L(\theta_1, \theta_2, ..., \theta_k)/\partial\theta_i$, $i = 1, 2, ..., k$ y resolviendo el sistema de ecuaciones resultante.

EJEMPLO 7-7

Sea que X tenga una distribución normal con media μ y varianza σ^2, con μ y σ^2 desconocidas. La función de verosimilitud de una muestra aleatoria de tamaño n es

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/(2\sigma^2)}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-(1/2\sigma^2)\sum_{i=1}^{n}(x_i-\mu)^2}$$

y

$$ln\, L(\mu, \sigma^2) = -\frac{n}{2} ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Entonces,

$$\frac{\partial\, ln\, L(\mu, \sigma^2)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\frac{\partial\, ln\, L(\mu, \sigma^2)}{\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

Las soluciones de la ecuación anterior producen los estimadores de máxima verosimilitud

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = X$$

y

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - X)^2$$

Propiedades del estimador de máxima verosimilitud

El método de máxima verosimilitud es con frecuencia el que prefieren los estadísticos matemáticos, pues por lo general es fácil de usar y produce estimadores con propiedades estadísticas buenas. A continuación se resumen estas propiedades:

Propiedades del estimador de máxima verosimilitud

Bajo condiciones muy generales y no restrictivas, cuando el tamaño de la muestra n es grande y si $\hat{\Theta}$ es el estimador de máxima verosimilitud del parámetro θ, entonces

1) $\hat{\Theta}$ es un estimador aproximadamente insesgado para $\theta\left[E(\hat{\Theta}) \simeq \theta\right]$,
2) la varianza de $\hat{\Theta}$ es casi tan pequeña como la varianza que se obtendría con cualquier otro estimador, y
3) $\hat{\Theta}$ tiene una distribución normal aproximada.

Las propiedades 1 y 2 establecen en esencia que el estimador de máxima verosimilitud es aproximadamente un MVUE. Este es un resultado muy conveniente y, aunado al hecho de que es bastante sencillo de obtener en muchas situaciones y tiene una distribución normal asintótica ("asintótica" quiere decir "cuando n es grande"), explica por qué la técnica de estimación de máxima verosimilitud tiene un uso tan generalizado. Para usar la estimación de máxima verosimilitud, recuérdese que la distribución de la población debe ser conocida o bien debe suponerse.

Para ilustrar la naturaleza tanto de "muestra grande" o asintótica de las propiedades anteriores, considérese el estimador de máxima verosimilitud de σ^2, la varianza de la distribución normal, en el ejemplo 7-7. Es fácil demostrar que

$$E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

El sesgo es

$$E(\hat{\sigma}^2) = \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = \frac{-\sigma^2}{n}$$

Puesto que el sesgo es negativo, $\hat{\sigma}^2$ tiende a subestimar el valor real de σ^2. Obsérvese que el sesgo tiende a cero cuando n se incrementa. Por lo tanto, $\hat{\sigma}^2$ es un estimador asintóticamente insesgado de σ^2.

Se presenta a continuación otra importante y útil propiedad de los estimadores de máxima verosimilitud.

La propiedad de invarianza

Sean $\hat{\Theta}_1, \hat{\Theta}_2, ..., \hat{\Theta}_k$ los estimadores de máxima verosimilitud de los parámetros $\theta_1, \theta_2, ..., \theta_k$. Entonces el estimador de máxima verosimilitud de cualquier función $h(\theta_1, \theta_2, ..., \theta_k)$ de estos parámetros es la misma función $h(\hat{\Theta}_1, \hat{\Theta}_2, ..., \hat{\Theta}_k)$ de los estimadores $\hat{\Theta}_1, \hat{\Theta}_2, ..., \hat{\Theta}_k$.

EJEMPLO 7-8

En el caso de la distribución normal, los estimadores de máxima verosimilitud de μ y σ^2 fueron $\hat{\mu} = \overline{X}$ y $\sigma^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/n$. Para obtener el estimador de máxima verosimilitud de la función $h(\mu, \sigma^2) = \sqrt{\sigma^2} = \sigma$, se sustituyen los estimadores $\hat{\mu}$ y $\hat{\sigma}^2$ en la función h, obteniéndose

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right]^{1/2}$$

Por tanto, el estimador de máxima verosimilitud de la desviación estándar σ *no* es la desviación estándar muestral S.

Complicaciones al usar la estimación de máxima verosimilitud

Aun cuando el método de máxima verosimilitud es una técnica excelente, en ocasiones surgen complicaciones en su uso. Por ejemplo, no siempre es sencillo maximizar la función de verosimilitud porque la(s) ecuación(es) obtenidas de $dL(\theta)/d\theta = 0$ puede(n) ser difícil(es) de resolver. Para un ejemplo de esto, remítase al ejercicio 7-19. Además, no siempre puede ser posible la utilización directa de los métodos del cálculo para determinar el máximo de $L(\theta)$. Esto se ilustra en el ejemplo siguiente.

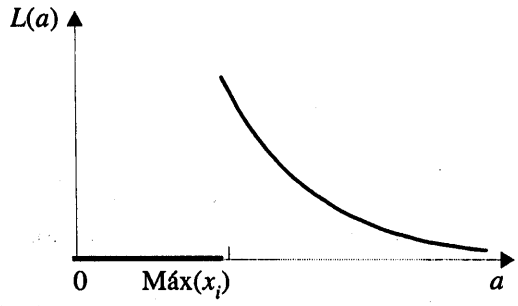


Figura 7-5 La función de verosimilitud de la distribución uniforme del ejemplo 7-9.

EJEMPLO 7-9

Sea que X tenga una distribución uniforme en el intervalo de 0 a a. Puesto que la función de densidad es $f(x) = 1/a$ para $0 \le x \le a$ y cero en caso contrario, la función de verosimilitud de una muestra aleatoria de tamaño n es

$$L(a) = \prod_{i=1}^{n} \frac{1}{a} = \frac{1}{a^n}$$

si $0 \le x_1 \le a$, $0 \le x_2 \le a$, ..., $0 \le x_n \le a$. Obsérvese que la pendiente de esta función no es cero en todos los puntos. Es decir, en tanto máx$(x_i) \le a$, la verosimilitud es $1/a^n$, que es positivo, pero cuando $a <$ máx(x_i), la verosimilitud cae hasta cero, como se ilustra en la figura 7-5. Por lo tanto, los métodos del cálculo no pueden usarse directamente debido a que el valor máximo de la función de verosimilitud ocurre en un punto de discontinuidad. Sin embargo, puesto que $(d/da)(a^{-n}) = -n/a^{n+1}$ es menor que cero para todos los valores de $a > 0$, a^{-n} es una función decreciente de a. Esto implica que el máximo de la función de verosimilitud $L(a)$ ocurre en el punto frontera inferior. La figura indica claramente que $L(a)$ podría maximizarse haciendo $\hat{a}$ igual al menor valor que podría asumir lógicamente, que es máx(x_i). Evidentemente, a no puede ser menor que la observación muestral más grande, por lo que es razonable hacer $\hat{a}$ igual al valor muestral más grande.

EJERCICIOS DE LA SECCIÓN 7-4

7-15. Considere la distribución de Poisson

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Encuentre el estimador de máxima verosimilitud de λ, con base en una muestra aleatoria de tamaño n.

7-16. Considere la distribución exponencial modificada

$$f(x) = \lambda e^{-\lambda(x-\theta)}, \quad x \ge \theta$$

Cuando $\theta = 0$, esta densidad se reduce a la distribución exponencial común. Cuando $\theta > 0$, sólo hay probabilidades positivas a la derecha de θ.

a) Encuentre el estimador de máxima verosimilitud de λ y θ, con base en una muestra aleatoria de tamaño n.

b) Describa una situación práctica en la que cabría esperar que la distribución exponencial modificada es un modelo recomendable.

7-17. Sea X una variable aleatoria geométrica con parámetro p. Encuentre el estimador de máxima verosimilitud de p, con base en una muestra aleatoria de tamaño n.

7-18. Sea X una variable aleatoria con la siguiente distribución de probabilidad:

$$f(x)=\begin{cases}(\alpha+1)x\alpha, & 0<x<1\\ 0 & , \text{ en caso contrario}\end{cases}$$

Encuentre el estimador de máxima verosimilitud de α, con base en una muestra aleatoria de tamaño n.

7-19. Considere la distribución de Weibull

$$f(x)=\begin{cases}\dfrac{\beta}{\delta}\left(\dfrac{x}{\delta}\right)^{\beta-1}e^{-\left(\frac{x}{\delta}\right)^{\beta}}, & 0<x\\ 0 & , \text{ en caso contrario}\end{cases}$$

a) Encuentre la función de verosimilitud con base en una muestra aleatoria de tamaño n. Encuentre la verosimilitud logarítmica.

b) Demuestre que la verosimilitud logarítmica se maximiza al resolver las ecuaciones

$$\beta=\left[\frac{\sum_{i=1}^{n}x_i^{\beta}\,ln(x_i)}{\sum_{i=1}^{n}x_i^{\beta}}-\frac{\sum_{i=1}^{n}ln(x_i)}{n}\right]^{-1}$$

$$\delta=\left[\frac{\sum_{i=1}^{n}x_i^{\beta}}{n}\right]^{1/\beta}$$

c) ¿Qué complicaciones se presentan al resolver las dos ecuaciones del inciso b)?

7-20. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de una distribución gamma con parámetros r y λ.

a) Encuentre la función de verosimilitud y la verosimilitud logarítmica.

b) Encuentre las ecuaciones que definen los estimadores de máxima verosimilitud de r y λ. ¿Pueden resolverse de manera explícita?

c) Demuestre que el estimador de máxima verosimilitud de $\mu = r/\lambda$ es $=$.

7-21. Sea que $X_1, X_2, ..., X_n$ tengan una distribución uniforme en el intervalo de 0 a a. Recuérdese que el estimador de máxima verosimilitud de a es $\hat{a} = \text{máx}(X_i)$.

a) Argumente de manera intuitiva por qué $\hat{a}$ no puede ser un estimador insesgado de a.

b) Suponga que $E(\hat{a}) = na/(n + 1)$. ¿Es razonable que $\hat{a}$ subestime consis tentemente a a? Demuestre que el sesgo en el estimador tiende a cero cuando n se incrementa.

c) Proponga un estimador insesgado de a.

d) Sea $Y = \text{máx}(X_i)$. Use el hecho de que $Y \le y$ si y sólo si toda $X_i \le y$ para derivar la función de distribución acumulada de Y. Demuestre después que la función de densidad de probabilidad de Y es

$$f(y)=\begin{cases}\dfrac{ny^{n-1}}{a^n}, & 0<y\le a\\ 0 & , \text{ en caso contrario}\end{cases}$$

Utilice este resultado para demostrar que el estimador de máxima verosimilitud de a es sesgado.

7-5 DISTRIBUCIONES DE MUESTREO

La inferencia estadística se refiere a la toma de **decisiones** acerca de una población con base en la información contenida en una muestra aleatoria de esa población. Por ejemplo, puede interesarnos el volumen de llenado medio de una lata de refresco. Se requiere que el volumen de llenado medio de la población sea 300 ml. Un ingeniero toma una muestra aleatoria de 25 latas y calcula el volumen de llenado promedio muestral como $\bar{x}$ = 298 ml. Probablemente el ingeniero decidirá que la media poblacional es μ = 300 ml, aun cuando la media muestral fue 298 ml

porque sabe que la media muestral es una estimación razonable de μ y que una media muestral de 298 ml tiene muchas posibilidades de ocurrir, incluso si la media poblacional real es $\mu = 300$ ml. De hecho, si la media real es 300 ml, entonces las pruebas repetidas de 25 latas, quizá cada cinco minutos, producirían valores de $\bar{x}$ que varíen tanto arriba como abajo de $\mu = 300$ ml.

La media muestral es un estadístico; es decir, es una variable aleatoria que depende de los resultados obtenidos en cada muestra particular. Puesto que un estadístico es una variable aleatoria, tiene una distribución de probabilidad.

Definición

A la distribución de probabilidad de un estadístico se le llama **distribución de muestreo**.

Por ejemplo, a la distribución de probabilidad de $\bar{X}$ se le llama la **distribución de muestreo de la media**.

La distribución de muestreo de un estadístico depende de la distribución de la población, del tamaño de la muestra y del método utilizado para seleccionar la muestra. En la siguiente sección de este capítulo se presenta la que quizá sea la distribución de muestreo más importante. Otras distribuciones de muestreo y sus aplicaciones se ilustrarán en detalle en los dos capítulos siguientes.

7-6 DISTRIBUCIONES DE MUESTREO DE MEDIAS

Considérese la determinación de la distribución de muestreo de la media muestral $\bar{X}$. Suponga que se toma una muestra aleatoria de tamaño n de una población normal con media μ y varianza σ^2. Entonces, cada observación de esta muestra, digamos, $X_1, X_2, ..., X_n$, es una variable aleatoria con una distribución normal e independiente, con media μ y varianza σ^2. Entonces, por la propiedad de reproducción de la distribución normal, ecuación 6-39 del capítulo 6, se concluye que la media muestral

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

tiene una distribución normal con media

$$\mu_{\bar{X}} = \frac{\mu + \mu + \cdots + \mu}{n} = \mu$$

y varianza

$$\sigma^2_{\bar{X}} = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma}{n^2} = \frac{\sigma^2}{n}$$

Si el muestreo se hace en una población que tiene una distribución de probabilidad desconocida, la distribución de muestreo de la media muestral seguirá siendo aproximadamente normal con media μ y varianza σ^2/n, si el tamaño de la muestra n es grande. Este es uno de los

teoremas más útiles de la estadística, llamado el **teorema del límite central**. El enunciado del mismo es el siguiente:

Teorema 7-2: El teorema del límite central

Si $X_1, X_2, ..., X_n$ es una muestra aleatoria de tamaño n tomada de una población (sea finita o infinita) con media μ y varianza finita σ^2, y si $\bar{X}$ es la media muestral, entonces la forma límite de la distribución de

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \tag{7-7}$$

cuando $n \to \infty$, es la distribución normal estándar.

La aproximación normal de $\bar{X}$ depende del tamaño de la muestra n. En la figura 7-6*a* se muestra la distribución obtenida para los lanzamientos de un solo dado equilibrado de seis caras. Las probabilidades son iguales (1/6) para todos los valores obtenidos, 1, 2, 3, 4, 5 o 6. En la figura 7-6*b* se muestra la distribución del resultado promedio que se obtiene cuando se lanzan dos dados, y en las figuras 7-6*c*, 7-6*d* y 7-6*e* se muestran las distribuciones de los resultados promedio que se obtienen cuando se lanzan tres, cinco y diez dados, respectivamente. Obsérvese que, si bien la población (un dado) está relativamente lejos de ser normal, la distribución de los promedios se aproxima de manera razonable por la distribución normal para tamaños de la muestra tan pequeños como cinco. (Sin embargo, las distribuciones del lanzamiento de los dados son discretas, en tanto que la distribución normal es continua.) Aun cuando el teorema del límite central funcionará bien para muestras pequeñas ($n = 4, 5$) en la mayoría de los casos, en particular cuando la población es continua, unimodal y simétrica, en otras situaciones se necesitarán muestras más grandes, dependiendo de la forma de la población. En muchos casos de interés práctico, si $n \geq 30$, la aproximación normal será satisfactoria independientemente de la forma de la población. Si $n < 30$, el teorema del límite central funcionará si la distribución de la población no se aparta significativamente de la distribución normal.

EJEMPLO 7-10

Una compañía de productos electrónicos fabrica resistores que tienen una resistencia media de 100 Ω y una desviación estándar de 10 Ω. La distribución de la resistencia es normal. Encuentre la probabilidad de que una muestra aleatoria de $n = 25$ resistores tenga una resistencia promedio menor que 95 Ω.

Obsérvese que la distribución de muestreo de $\bar{X}$ es normal, con media $\mu_{\bar{X}} = 100$ Ω y una desviación estándar de

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$

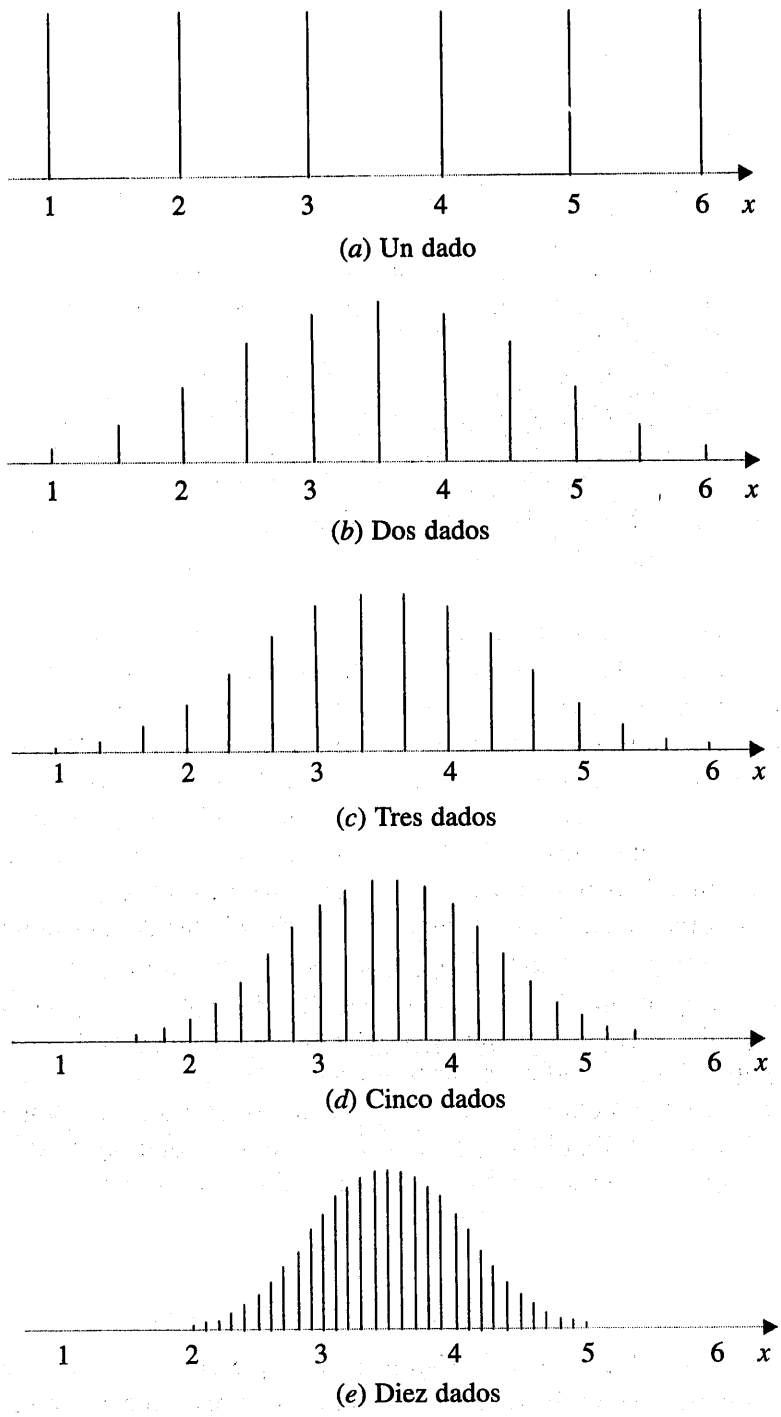


Figura 7-6 Distribuciones de los resultados promedio para el lanzamiento de dados. [Adaptado con permiso de Box, Hunter y Hunter.]

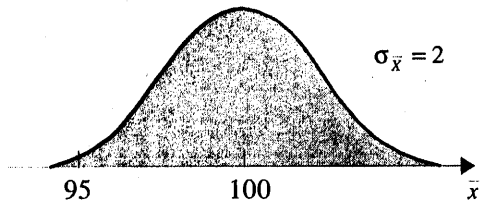


Figura 7-7 Probabilidad del ejemplo 7-10.

Por lo tanto, la probabilidad pedida corresponde con el área sombreada de la figura 7-7. Al estandarizar el punto $\bar{X} = 95$ de la figura 7-7, se encuentra que

$$z = \frac{95 - 100}{2} = -2.5$$

y por lo tanto

$$P(\bar{X} < 95) = P(Z < -2.5)$$
$$= 0.0062$$

En el ejemplo siguiente se aplica el teorema del límite central.

EJEMPLO 7-11

Suponga que una variable aleatoria X tiene una distribución uniforme continua

$$f(x) = \begin{cases} 1/2, & 4 \le x \le 6 \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de la media muestral de una muestra aleatoria de tamaño $n = 40$.

La media y la varianza de X son $\mu = 5$ y $\sigma^2 = (6-4)^2/12 = 1/3$. El teorema del límite central indica que la distribución de $\bar{X}$ es aproximadamente normal con media $\mu_{\bar{X}} = 5$ y varianza $\sigma^2_{\bar{x}} = \sigma^2/n = 1/[3(40)] = 1/120$. Las distribuciones de X y $\bar{X}$ se muestran en la figura 7-8.

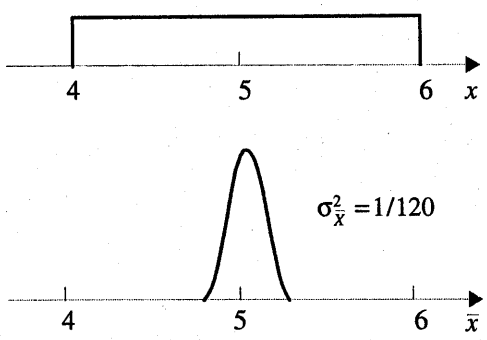


Figura 7-8 Las distribuciones de X y $\bar{X}$ del ejemplo 7-11.

Considérese ahora el caso en que se tienen dos poblaciones independientes. Sea que la primera población tenga media μ_1 y varianza σ_1^2 y la segunda población tenga media μ_2 y varianza σ_2^2. Suponga que ambas poblaciones tienen una distribución normal. Entonces, utilizando el hecho de que las combinaciones lineales de variables aleatorias normales independientes siguen una distribución normal (véase la ecuación 6-39), puede decirse que la distribución de muestreo de $\bar{X}_1 - \bar{X}_2$ es normal con media

$$\begin{aligned} \mu_{\bar{X}_1-\bar{X}_2} &= \mu_{\bar{X}_1} - \mu_{\bar{X}_2} \\ &= \mu_1 - \mu_2 \end{aligned} \tag{7-8}$$

y la varianza

$$\begin{aligned} \sigma^2_{\bar{X}_1-\bar{X}_2} &= \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2} \\ &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \end{aligned} \tag{7-9}$$

Si las dos poblaciones no tienen una distribución normal, pero si ambos tamaños de las muestras n_1 y n_2 son mayores que 30, puede aplicarse el teorema del límite central y suponer que $\bar{X}_1$ y $\bar{X}_2$ siguen aproximadamente una distribución normal independiente. Por lo tanto, la distribución de muestreo de $\bar{X}_1 - \bar{X}_2$ es aproximadamente normal con media y varianza dadas por las ecuaciones 7-8 y 7-9, respectivamente. Si n_1 o n_2 es menor que 30, entonces la distribución de muestreo de $\bar{X}_1 - \bar{X}_2$ seguirá siendo aproximadamente normal con media y varianza dadas por las ecuaciones 7-8 y 7-9, siempre que la población de la que se tomó la muestra pequeña no sea marcadamente diferente de la normal. Esto puede resumirse con la siguiente definición:

Definición

Si se tienen dos poblaciones independientes con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2, y si $\bar{X}_1$ y $\bar{X}_2$ son las medias muestrales de dos variables aleatorias independientes de tamaños n_1 y n_2 de estas poblaciones, entonces la distribución de muestreo de

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}} \tag{7-10}$$

es aproximadamente normal estándar, si se cumplen las condiciones del teorema del límite central. Si las dos poblaciones son normales, entonces la distribución de muestreo de Z es exactamente normal estándar.

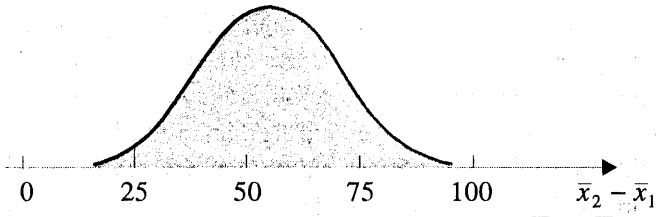


Figura 7-9 La distribución de muestreo de $\bar{X}_2 - \bar{X}_1$ del ejemplo 7-12.

EJEMPLO 7-12

La vida efectiva de un componente usado en la turbina de un avión es una variable aleatoria con una media de 5 000 hr y una desviación estándar de 40 hr. La distribución de la vida efectiva se acerca bastante a una distribución normal. El fabricante de las turbinas introduce una mejora en el proceso de manufactura de este componente, la cual incrementa la vida media a 5 050 hr y disminuye la desviación estándar a 30 hr. Suponga que se selecciona una muestra aleatoria de $n_1 = 16$ componentes del proceso "antiguo" y que se selecciona una muestra aleatoria de $n_2 = 25$ componentes del proceso "mejorado". ¿Cuál es la probabilidad de que la diferencia en las medias de las dos muestras $\bar{X}_2 - \bar{X}_1$ sea al menos 25 hr? Suponga que el proceso antiguo y el mejorado pueden considerarse poblaciones independientes.

Para resolver este problema, se observa primero que la distribución de $\bar{X}_1$ es normal con media $\mu_1 = 5\ 000$ hr y desviación estándar $\sigma_1/\sqrt{n_1} = 40/\sqrt{16} = 10$ hr, y que la distribución de $\bar{X}_2$ es normal con media $\mu_2 = 5050$ hr y desviación estándar $\sigma_2/\sqrt{n_2} = 30/\sqrt{25} = 6$ hr. Entonces la distribución de $\bar{X}_2 - \bar{X}_1$ es normal con media $\mu_2 - \mu_1 = 5\ 050 - 5\ 000 = 50$ hr, y varianza $\sigma_2^2/n_2 + \sigma_1^2/n_1 = (6)^2 + (10)^2 = 136$ hr^2. Esta distribución de muestreo se ilustra en la figura 7-9. La probabilidad de que $\bar{X}_2 - \bar{X}_1 \geq 25$ es el área sombreada de la distribución normal en esta figura.

Para el valor que corresponde a $\bar{x}_2 - \bar{x}_1 = 25$ en la figura 7-9, se encuentra

$$z = \frac{25 - 50}{\sqrt{136}} = -2.14$$

y se encuentra que

$$P(\bar{X}_2 - \bar{X}_1 \geq 25) = P(Z \geq -2.14)$$
$$= 0.9838$$

EJERCICIOS DE LA SECCIÓN 7-6

7-22. Se fabrica tubería PVC con un diámetro medio de 1.01 pulgadas y una desviación estándar de 0.003 pulgadas. Encuentre la probabilidad de que una muestra aleatoria de $n = 9$ secciones de tubería tenga un diámetro muestral medio mayor que 1.009 pulgadas y menor que 1.012 pulgadas.

7-23. Suponga que se seleccionan al azar muestras de tamaño $n = 25$ de una población normal con media 100 y desviación estándar 10. ¿Cuál es la probabilidad de que la media muestral esté en el intervalo de $\mu_{\bar{X}} - 1.8\,\sigma_{\bar{X}}$ a $\mu_{\bar{X}} + 1.0\,\sigma_{\bar{X}}$?

7-24. Una fibra sintética usada en la fabricación de alfombras tiene una resistencia a la tensión distribuida normalmente con una media de 75.5 psi y una desviación estándar de 3.5 psi. Encuentre la probabilidad de que una muetra aleatoria de $n = 6$ fibras tenga una resistencia a la tensión media muestral que exceda 75.75 psi.

7-25. Considere la fibra sintética del ejercicio anterior. ¿Cómo cambia la desviación estándar de la media muestral cuando el tamaño de la muestra se incrementa de $n = 6$ a $n = 49$?

7-26. La fuerza de compresión del concreto tiene una media de 2 500 psi y una desviación estándar de 50 psi. Encuentre la probabilidad de que una muestra aleatoria de $n = 5$ observaciones tenga una fuerza de compresión media muestral que esté en el intervalo de 2 499 psi a 2 510 psi.

7-27. Considere las observaciones del concreto del ejercicio anterior. ¿Cuál es el error estándar de la media muestral?

7-28. Una población normal tiene media 100 y varianza 25. ¿De qué tamaño deberá ser la muestra aleatoria si se quiere que el error estándar del promedio muestral sea 1.5?

7-29. Suponga que la variable aleatoria X tiene la distribución continua uniforme

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{en caso contrario} \end{cases}$$

Suponga que se selecciona una muestra aleatoria de $n = 12$ observaciones de esta distribución. ¿Cuál es la distribución de probabilidad de $\bar{X} - 6$? Encuentre la media y la varianza de esta cantidad.

7-30. Suponga que X tiene la distribución uniforme discreta

$$f(x) = \begin{cases} {}^{1}/_{3}, & x = 1, 2, 3 \\ 0, & \text{en caso contrario} \end{cases}$$

Se selecciona una muestra aleatoria de $n = 36$ de esta población. Encuentre la probabilidad de que la media muestral sea mayor que 2.1 pero menor que 2.5, suponiendo que la media muestral se mide hasta la décima más próxima.

7-31. La cantidad de tiempo que un pasajero pasa esperando en el mostrador de registro de un aeropuerto es una variable aleatoria con media 8.2 minutos y desviación estándar 1.5 minutos. Suponga que se observa una muestra aleatoria de $n = 49$ pasajeros. Encuentre la probabilidad de que el tiempo promedio de espera en la fila de estos pasajeros sea:

a) Menor que 10 minutos.
b) Entre 5 y 10 minutos.
c) Menor que 6 minutos.

7-32. Se selecciona una muestra aleatoria de tamaño $n_1 = 16$ de una población normal con una media de 75 y una desviación estándar de 8. Se toma una segunda muestra aleatoria de tamaño $n_2 = 9$ de otra población normal con media 70 y desviación estándar 12. Sean $\bar{X}_1$ y $\bar{X}_2$ las dos medias muestrales. Encuentre:

a) La probabilidad de que $\bar{X}_1 - \bar{X}_2$ exceda 4.
b) La probabilidad de que $3.5 \le \bar{X}_1 - \bar{X}_2 \le 5.5$.

7-33. Una compañía de insumos electrónicos está comparando la brillantez de dos tipos diferentes de cinescopio para emplearlos en sus televisores. El cinescopio tipo A tiene una brillantez media de 100 y una desviación estándar de 16, en tanto que el cinescopio tipo B tiene una brillantez media desconocida, pero la desviación estándar se supone idéntica a la del tipo A. Se selecciona una muestra aleatoria de $n = 25$ cinescopios de cada tipo y se calcula $\bar{X}_B - \bar{X}_A$. Si μ_B es igual o excede a μ_A, el fabricante preferiría adoptar el cinescopio tipo B. La diferencia observada es $\bar{x}_B - \bar{x}_A = 3.5$. ¿Qué decisión tomaría usted, y por qué?

7-34. La elasticidad de un polímero es afectada por la concentración del reactivo. Cuando se usa una concentración baja, la elasticidad media real es 55, y cuando se usa una concentración alta la elasticidad media es 60. La desviación estándar de la elasticidad es 4, independientemente de la concentración. Si se toman dos muestras aleatorias de tamaño 16, encuentre la probabilidad de que $\bar{X}_{\text{alta}} - \bar{X}_{\text{baja}} \ge 2$.

7-7 INTRODUCCIÓN A LOS INTERVALOS DE CONFIANZA

En muchas situaciones, la estimación puntual de un parámetro no le ofrece información completa a un ingeniero. Por ejemplo, considérese el problema de la conductividad térmica del ejemplo 7-2. La estimación puntual de la conductividad térmica para este material particular fue $\bar{x}$ = 41.924 Btu/hr-pie-°F. Es poco probable que la conductividad térmica media real μ sea exactamente igual a este valor, lo cual plantea la pregunta, ¿qué tan cerca está $\bar{x}$ de la media real? Como se señaló en la sección 7-3.3, el error estándar de la estimación puntual es una guía aproximada de la precisión de una estimación. Otro enfoque es usar un **intervalo de confianza** para expresar el grado de incertidumbre asociado con una estimación puntual.

La estimación del intervalo de confianza de un parámetro desconocido θ es un intervalo de la forma $l \leq \theta \leq u$, donde los puntos finales l y u dependen del valor numérico del estadístico muestral $\hat{\Theta}$ para una muestra particular. Puesto que muestras diferentes producirán valores diferentes de $\hat{\theta}$ y, en consecuencia, valores diferentes de los puntos finales l y u, estos puntos finales son valores de variables aleatorias, digamos, L y U, respectivamente. A partir de la distribución de muestreo de la media del estadístico $\hat{\Theta}$ podrán determinarse los valores de L y U tales que el siguiente enunciado de probabilidad sea verdadero:

$$P(L \leq \theta \leq U) = 1 - \alpha \tag{7-11}$$

donde $0 < \alpha < 1$. Por tanto, se tiene una probabilidad de $1 - \alpha$ de seleccionar una muestra que producirá un intervalo que incluya al valor verdadero de $\hat{\theta}$.

Al intervalo que resulta

$$l \leq \theta \leq u \tag{7-12}$$

se le llama un **intervalo de confianza** del $100(1 - \alpha)$ por ciento para el parámetro θ. A las cantidades l y u se les llama **límites de confianza inferior y superior**, respectivamente, y a $1 - \alpha$ se le llama el **coeficiente de confianza**. La interpretación de un intervalo de confianza es que, si se toma un número infinito de muestras aleatorias y se calcula un intervalo de confianza del $100(1 - \alpha)$ por ciento para θ en cada muestra, entonces $100(1 - \alpha)$ por ciento de estos intervalos incluirán el valor real de θ.

La situación se ilustra en la figura 7-10, donde se presentan varios intervalos de confianza del $100(1 - \alpha)$ por ciento para el parámetro θ de una distribución. Los puntos en el centro de los intervalos indican la estimación puntual de θ (es decir, $\hat{\theta}$). Obsérvese que uno de los 15 intervalos no incluye al valor real de θ. Si éste fuera un intervalo de confianza de 95%, a la larga sólo 5% de los intervalos no incluirían a θ.

En la práctica, se obtiene una sola muestra aleatoria y se calcula un solo intervalo de confianza. Puesto que este intervalo incluirá o bien no incluirá al valor verdadero de θ, no es razonable asignar un nivel de probabilidad a este evento específico. El planteamiento apropiado es que el intervalo observado $[l, u]$ comprende al valor verdadero de θ con una confianza de $100(1 - \alpha)$. Este enunciado tiene una interpretación de frecuencia; es decir, no se sabe si el enunciado es verdadero para esta muestra específica, pero el método usado para obtener el intervalo $[l, u]$ produce enunciados verdaderos $100(1 - \alpha)$ por ciento de las veces.

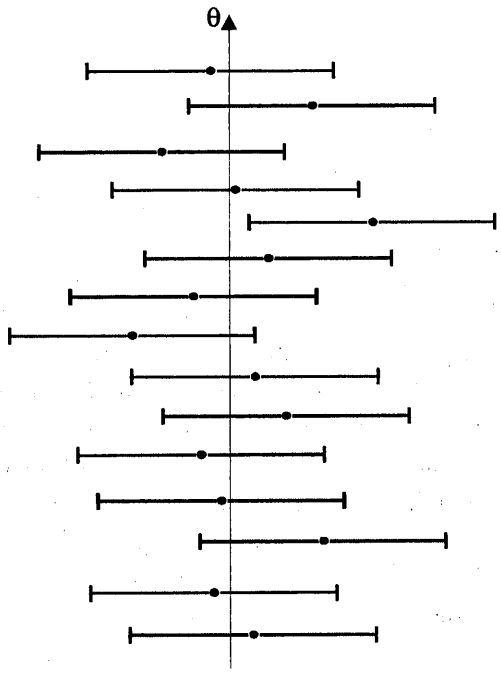


Figura 7-10 Construcción repetida de un intervalo de confianza para θ.

La longitud $\theta - l$ del intervalo de confianza observado es una medida importante de la calidad de la información que se obtiene de la muestra. A la longitud del semiintervalo $\theta - l$ o $u - \theta$ se le llama la **precisión** del estimador. Entre más largo sea el intervalo de confianza, se tendrá una confianza mayor de que el intervalo en realidad contiene el valor verdadero de θ. Por otra parte, entre más largo sea el intervalo, se tendrá menos información acerca del valor verdadero de θ. En una situación ideal, se obtiene un intervalo relativamente corto con alta confianza.

En muchas situaciones prácticas es sencillo encontrar los puntos extremos que definen el intervalo de confianza para un parámetro. Por ejemplo, los puntos extremos del intervalo de confianza para la media μ de una distribución normal incluyen el error estándar de la media muestral $\bar{X}$. De hecho, el intervalo de confianza de μ se encuentra sumando y restando a la media muestral un múltiplo del error estándar $\sigma/\sqrt{n}$ o del error estándar estimado $\varsigma/\sqrt{n}$. Este método se usó en el ejemplo 7-2 para encontrar un intervalo en el que se tuviera una gran confianza de que incluiría la conductividad térmica media real del hierro Armco. En los capítulos 8 y 9 se presentan los detalles de la construcción de intervalos de confianza en diversas situaciones prácticas.

Los intervalos de confianza guardan una estrecha relación con otra importante técnica estadística para la toma de decisiones llamada **prueba de hipótesis**. Las hipótesis son tan sólo enunciados acerca de los parámetros de las distribuciones de probabilidad. El objetivo es tomar decisiones acerca de estos enunciados. Con frecuencia, estas decisiones pueden tomarse examinando el rango de valores razonables del parámetro en un intervalo de confianza. En los capítulos 8 y 9 se revisan e ilustran diferentes escenarios de prueba de hipótesis. En estos dos capítulos se integran la prueba de hipótesis y los intervalos de confianza porque prácticamente todos los paquetes de software los presentan así.

Ejercicios complementarios

7-35. Suponga que una variable aleatoria tiene una distribución normal con media μ y desviación estándar σ^2, y se extrae una muestra aleatoria de cinco observaciones de esta distribución. ¿Cuál es la función de densidad de probabilidad conjunta de la muestra?

7-36. Los transistores tienen una vida útil que sigue una distribución exponencial con parámetro λ. Se toma una muestra aleatoria de n transistores. ¿Cuál es la función de densidad de probabilidad conjunta de la muestra?

7-37. Suponga que X tiene una distribución uniforme en el intervalo de 0 a 1. Considere una muestra aleatoria de X de tamaño 4. ¿Cuál es la función de densidad de probabilidad conjunta de la muestra?

7-38. Un especialista en compras ha adquirido 25 resistores del vendedor 1 y 30 resistores del vendedor 2. Sea que $X_{1,1}, X_{1,2}, ..., X_{1,25}$ represente las resistencias observadas del vendedor 1, las cuales se suponen con una distribución normal e independiente con media 100 Ω y desviación estándar 1.5 Ω. De manera similar, sea que $X_{2,1}, X_{2,2}, ..., X_{2,30}$ represente las resistencias observadas del vendedor 2, las cuales se suponen con una distribución normal e independiente con media 105 Ω y desviación estándar 2.0 Ω. ¿Cuál es la distribución de muestreo de $\overline{X}_1 - \overline{X}_2$?

7-39. Considere el problema de los resistores del ejercicio 7-38. ¿Cuál es el error estándar de $\overline{X}_1 - \overline{X}_2$?

7-40. Se toma una muestra aleatoria de 36 observaciones de una distribución normal con media 50 y desviación estándar 12. Encuentre la probabilidad de que la media muestral esté en el intervalo $47 \le \overline{X} \le 53$.

7-41. ¿El supuesto de normalidad es importante en el ejercicio 7-40? ¿Por qué?

7-42. Se prueba una muestra aleatoria de $n = 9$ elementos estructurales para la resistencia a la compresión. Se sabe que la resistencia a la compresión media verdadera es $\mu = 5\ 500$ psi y que la desviación estándar es $\sigma = 100$ psi. Encuentre la probabilidad de que la media muestral de la resistencia a la compresión exceda 4 985 psi.

7-43. Una población normal tiene una media conocida de 50 y una varianza conocida de $\sigma^2 = 2$. Se selecciona una muestra aleatoria de $n = 16$ de esta población, y la media muestral es $\bar{x} = 52$. ¿Qué tan fuera de lo común es este resultado?

7-44. Se toma una muestra aleatoria de tamaño $n = 16$ de una población normal con $\mu = 40$ y $\sigma^2 = 5$. Encuentre la probabilidad de que la media muestral sea menor o igual que 37.

7-45. Un fabricante de dispositivos semiconductores toma una muestra aleatoria de 100 chips y los prueba, clasificando cada chip como defectuoso o no defectuoso. Sea $X_i = 0$ si el chip es no defectuoso y sea $X_i = 1$ si el chip está defectuoso. La proporción defectuosa de la muestra es

$$\hat{P} = \frac{X_1 + X_2 + \cdots + X_{100}}{100}$$

¿Cuál es la distribución de muestreo de la variable aleatoria $\hat{P}$?

7-46. Sea X una variable aleatoria con media μ y varianza σ^2. Dadas dos muestras aleatorias independientes de tamaños n_1 y n_2, con medias muestrales $\overline{X}_1$ y $\overline{X}_2$, demuestre que

$$\overline{X} = a\overline{X}_1 + (1-a)\overline{X}_2, \quad 0 < a < 1$$

es un estimador insesgado de μ. Si $\overline{X}_1$ y $\overline{X}_2$ son independientes, encuentre el valor de a que minimice el error estándar de $\overline{X}$.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

7-47. Un lote se compone de N transistores, y de éstos M ($M \leq N$) están defectuosos. Se seleccionan dos transistores al azar sin reemplazo de este lote y se determina si están defectuosos o no. La variable aleatoria

$$X_i = \begin{cases} 1, & \text{si el transistor } i \text{ no está defectuoso} \\ 0, & \text{si el transistor } i \text{ está defectuoso} \end{cases} \quad i = 1, 2$$

Determine la función de probabilidad conjunta de X_1 y X_2. ¿Cuáles son las funciones de probabilidad marginal de X_1 y X_2? ¿Las variables X_1 y X_2 son variables aleatorias independientes?

7-48. Cuando la desviación estándar muestral se basa en una muestra aleatoria de tamaño n de una población normal, es posible demostrar que S es un estimador sesgado de σ. Específicamente

$$E(S) = \sigma\sqrt{2/(n-1)}\ \Gamma(n/2)/\Gamma[(n-2)/2]$$

a) Use este resultado para obtener un estimador insesgado de σ de la forma $c_n S$, cuando la constante c_n depende del tamaño de la muestra n.

b) Encuentre el valor de c_n para $n = 10$ y $n = 25$. En general, ¿qué tan bueno es el desempeño de S como estimador de σ para n grande con respecto al sesgo?

7-49. Un operador mide dos veces una colección de n piezas seleccionadas al azar utilizando una regla graduada. Sea que X_i y Y_i denoten los valores medidos de la pieza i. Suponga que estas dos variables aleatorias son independientes, que tienen una distribución normal y que ambas tienen media verdadera μ_i y varianza σ^2.

a) Demuestre que el estimador de máxima verosimilitud de σ^2 es

$$\hat{\sigma}^2 = (1/4n)\sum_{i=1}^{n}(X_i - Y_i)^2$$

b) Demuestre que $\hat{\sigma}^2$ es un estimador sesgado de σ^2. ¿Qué ocurre con el sesgo cuando n se hace grande?

c) Encuentre un estimador insesgado de σ^2.

7-50. Estimador consistente. Otra forma de medir la cercanía de un estimador $\hat{\theta}$ del parámetro θ es en términos de consistencia. Si $\hat{\theta}_n$ es un estimador de θ basado en una muestra aleatoria de n observaciones, entonces $\hat{\theta}_n$ es consistente para θ si

$$\lim_{n\to\infty} P(|\hat{\theta} - \theta| < \epsilon) = 1$$

Por tanto, la consistencia es una propiedad para muestras grandes, la cual describe el comportamiento de $\hat{\theta}_n$ en el límite cuando n tiende a infinito. Suele ser difícil establecer la consistencia utilizando la definición anterior, aunque esto puede hacerse aplicando otros enfoques. Para ilustrar, demuestre que $\overline{X}$ es un estimador consistente de μ (cuando $\sigma^2 < \infty$) utilizando la desigualdad de Chebyshev (véase la sección 6-7).

7-51. Estadístico de orden. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de X, una variable aleatoria que tiene función de distribución $F(x)$. Al clasificar los elementos en orden de magnitud numérica creciente, se obtiene $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, donde $X_{(1)}$ es el elemento muestral más pequeño ($X_{(1)} = \text{mín}\{X_1, X_2, \ldots, X_n\}$) y $X_{(n)}$ es el elemento muestral más grande ($X_{(n)} = \text{máx}\{X_1, X_2, \ldots, X_n\}$). A $X_{(i)}$ se le llama el estadístico de orden i. Con frecuencia es de interés la distribución de alguno de los estadísticos de orden, en particular de los valores muestrales mínimo y máximo, $X_{(1)}$ y $X_{(n)}$, respectivamente. Demuestre que las funciones de distribución

EJERCICIOS PARA DESARROLLAR EL INTELECTO

acumulada de estos dos estadísticos de orden, denotadas respectivamente por $F_{X_{(1)}}(t)$ y $F_{X_{(n)}}(t)$ son

$$F_{X_{(1)}}(t) = 1 - [1 - F(t)]^n$$

$$F_{X_{(n)}}(t) = [F(t)]^n$$

Demuestre que si X es continua con función de densidad de probabilidad $f(x)$, entonces las distribuciones de probabilidad de $X_{(1)}$ y $X_{(n)}$ son

$$f_{X_{(1)}}(t) = n[1 - F(t)]^{n-1} f(t)$$

$$f_{X_{(n)}}(t) = n[F(t)]^{n-1} f(t)$$

7-52. Continuación del ejercicio 7-51. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una variable aleatoria de Bernoulli con parámetro p. Demuestre que

$$P(X_{(n)} = 1) = 1 - (1 - p)^n$$

$$P(X_{(1)} = 0) = 1 - p^n$$

Use los resultados del ejercicio 7-51.

7-53. Continuación del ejercicio 7-51. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una variable aleatoria normal con media μ y varianza σ^2. Utilizando los resultados del ejercicio 7-51, deduzca las funciones de densidad de probabilidad de $X_{(1)}$ y $X_{(n)}$.

7-54. Continuación del ejercicio 7-51. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una variable aleatoria exponencial con parámetro λ. Deduzca las funciones de distribución acumulada y las funciones de densidad de probabilidad de $X_{(1)}$ y $X_{(n)}$. Use el resultado del ejercicio 7-51.

7-55. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una variable aleatoria continua con función de distribución acumulada $F(x)$. Encuentre

$$E[F(X_{(n)})]$$

y

$$E[F(X_{(1)})]$$

7-56. Sea X una variable aleatoria con media μ y varianza σ^2, y sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de X de tamaño n. Demuestre que el estadístico $V = k\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2$ es un estimador insesgado de σ^2 para una elección apropiada de la constante k. Encuentre este valor de k.

8 Inferencia estadística para una sola muestra

PLAN GENERAL DEL CAPÍTULO

8-1 PRUEBA DE HIPÓTESIS
- 8-1.1 Hipótesis estadísticas
- 8-1.2 Pruebas de hipótesis estadísticas
- 8-1.3 Hipótesis de una y dos colas
- 8-1.4 Procedimiento general de las pruebas de hipótesis

8-2 INFERENCIA SOBRE LA MEDIA DE UNA POBLACIÓN, VARIANZA CONOCIDA
- 8-2.1 Pruebas de hipótesis sobre la media
- 8-2.2 Valores *P* en las pruebas de hipótesis
- 8-2.3 Error tipo II y elección del tamaño de la muestra
- 8-2.4 Prueba para muestras grandes
- 8-2.5 Algunos comentarios prácticos sobre las pruebas de hipótesis
- 8-2.6 Intervalo de confianza para la media
- 8-2.7 Método general para deducir un intervalo de confianza
- 8-2.8 Intervalos de confianza *bootstrap* (cargador inicial)

8-3 INFERENCIA SOBRE LA MEDIA DE UNA POBLACIÓN, VARIANZA DESCONOCIDA
- 8-3.1 Pruebas de hipótesis sobre la media
- 8-3.2 Valor *P* para una prueba *t*
- 8-3.3 Solución de computadora
- 8-3.4 Elección del tamaño de la muestra
- 8-3.5 Intervalo de confianza sobre la media

8-4 INFERENCIA SOBRE LA VARIANZA DE UNA POBLACIÓN NORMAL
- 8-4.1 Pruebas de hipótesis sobre la varianza de una población normal
- 8-4.2 Error β y elección del tamaño de la muestra
- 8-4.3 Intervalo de confianza sobre la varianza de una población normal

8-5 INFERENCIA SOBRE UNA PROPORCIÓN DE UNA POBLACIÓN
- 8-5.1 Pruebas de hipótesis sobre una proporción binomial
- 8-5.2 Error tipo II y elección del tamaño de la muestra
- 8-5.3 Intervalo de confianza sobre una proporción binomial

8-6 TABLA DE RESUMEN PARA PROCEDIMIENTOS DE INFERENCIA PARA UNA SOLA MUESTRA

8-7 PRUEBA DE LA BONDAD DEL AJUSTE

8-8 PRUEBAS DE TABLAS DE CONTINGENCIA

8-1 PRUEBA DE HIPÓTESIS

8-1.1 Hipótesis estadísticas

En el capítulo anterior se explicó cómo estimar un parámetro a partir de datos muestrales. Sin embargo, en muchos problemas de ingeniería es necesario decidir si se acepta o se rechaza un enunciado acerca de algún parámetro. Al enunciado se le llama **hipótesis**, y al procedimiento para tomar decisiones acerca de la hipótesis se le llama **prueba de hipótesis**. Se trata de uno de los aspectos más útiles de la inferencia estadística, ya que muchos problemas de toma de decisiones, pruebas o experimentos en el mundo de la ingeniería pueden formularse como problemas de prueba de hipótesis. Los autores del libro nos inclinamos por considerar la prueba estadística de hipótesis como la etapa de análisis de datos de un **experimento comparativo**, en el cual al ingeniero le interesa comparar, por ejemplo, la media de una población con un valor dado. Estos experimentos comparativos simples se encuentran con frecuencia en la práctica y proporcionan bases sólidas para los problemas de diseño experimental de mayor complejidad que se estudian en los capítulos 12 y 13. En este capítulo se discuten los experimentos comparativos en los que interviene una sola población, y la atención se centra en la prueba de hipótesis referentes a los parámetros de la población.

Se presenta ahora una definición formal de hipótesis estadística.

Definición

Una **hipótesis estadística** es un enunciado acerca de los parámetros de una o más poblaciones.

Puesto que se han venido usando distribuciones de probabilidad para representar poblaciones, una hipótesis estadística también podría considerarse como un enunciado acerca de la distribución de probabilidad de una variable aleatoria. En la hipótesis, por lo general interviene uno o más parámetros de esta distribución.

Por ejemplo, suponga que se tiene interés en la rapidez de combustión del propulsor sólido utilizado para impulsar los sistemas de expulsión de la tripulación de un avión. Entonces, la rapidez de combustión es una variable aleatoria que puede describirse con una distribución de probabilidad. Suponga que el interés se enfoca en la media de la rapidez de combustión (uno de los parámetros de esta distribución). Específicamente, quiere decidirse si la media de la rapidez de combustión es 50 cm/s o no. Esto puede expresarse formalmente como

$$\begin{aligned} H_0&: \mu = 50 \text{ cm/s} \\ H_1&: \mu \neq 50 \text{ cm/s} \end{aligned} \tag{8-1}$$

Al enunciado H_0: $\mu = 50$ cm/s de la ecuación 8-1 se le llama la **hipótesis nula**, y al enunciado H_1: $\mu \neq 50$ cm/s se le llama la **hipótesis alternativa**. Puesto que la hipótesis alternativa especifica el valor de μ que podría ser mayor o bien menor que 50 cm/s, se le llama **hipótesis**

alternativa de dos colas. En algunas situaciones, querría formularse una **hipótesis alternativa de una cola**, como en

$$H_0 : \mu = 50 \text{ cm/s} \qquad H_0 : \mu = 50 \text{ cm/s}$$

$$H_1 : \mu < 50 \text{ cm/s} \qquad H_1 : \mu > 50 \text{ cm/s} \tag{8-2}$$

Es importante recordar que las hipótesis son siempre enunciados acerca de la población o la distribución bajo estudio, no enunciados acerca de la muestra. El valor del parámetro poblacional especificado en la hipótesis nula (50 cm/s en el ejemplo anterior) se determina por lo general en una de tres formas. Primera, puede ser el resultado de la experiencia pasada, del conocimiento del proceso o, incluso, de pruebas o experimentos anteriores. Así, el objetivo de la prueba de hipótesis suele ser determinar si el valor del parámetro ha sufrido cambios. Segunda, este valor puede determinarse a partir de una teoría o modelo del proceso bajo estudio. En este caso, el objetivo de la prueba de hipótesis es verificar la teoría o el modelo. Una tercera situación surge cuando el valor del parámetro poblacional es resultado de consideraciones externas, como especificaciones de diseño o de ingeniería, o de obligaciones contractuales. En este caso, por lo general el objetivo de la prueba de hipótesis es una prueba de conformidad.

A un proceso que lleva a una decisión acerca de una hipótesis particular se le llama **prueba de una hipótesis**. Los procedimientos de prueba de hipótesis se basan en el uso de la información contenida en una muestra aleatoria de la población de interés. Si esta información es consistente con la hipótesis, entonces se concluirá que la hipótesis es verdadera; sin embargo, si esta información no es consistente con la hipótesis, se concluirá que la hipótesis es falsa. Se hace hincapié en que el carácter verdadero o falso de una hipótesis particular nunca puede conocerse con toda certeza, a menos que pueda examinarse la población completa, lo cual suele ser imposible en la mayoría de las situaciones prácticas. Por lo tanto, el procedimiento para la prueba de hipótesis deberá realizarse teniendo en mente la probabilidad de llegar a una conclusión incorrecta.

La estructura de los problemas de prueba de hipótesis es idéntica en todas las aplicaciones que se consideran en este texto. La hipótesis nula es la hipótesis que quiere probarse. El rechazo de la hipótesis nula lleva siempre a aceptar la hipótesis alternativa. En el tratamiento que se aplica aquí para la prueba de hipótesis, la hipótesis nula se enunciará especificando siempre un valor exacto del parámetro (como en el enunciado H_0: $\mu = 50$ cm/s de la ecuación 8-1). En la hipótesis alternativa siempre se permitirá que el parámetro tome varios valores (como en el enunciado H_1: $\mu \neq 50$ cm/s de la ecuación 8-1). Probar la hipótesis implica tomar una muestra aleatoria, calcular un **estadístico de prueba** de los datos muestrales, y utilizar después el estadístico de prueba para tomar una decisión acerca de la hipótesis nula.

8-1.2 Pruebas de hipótesis estadísticas

Para ilustrar los conceptos generales, considérese el problema de la rapidez de combustión de la carga propulsora que se introdujo antes. La hipótesis nula es que la rapidez de combustión media es 50 cm/s, y la alterna es que no es igual a 50 cm/s. Es decir, quiere probarse

$$H_0 : \mu = 50 \text{ cm/s}$$

$$H_1 : \mu \neq 50 \text{ cm/s}$$

Suponga que se prueba una muestra de $n = 10$ observaciones y se obtiene la media muestral $\bar{x}$ de la rapidez de combustión. La media muestral es una estimación de la verdadera media poblacional μ. Un valor de la media muestral $\bar{x}$ que está cerca del valor hipotético $\mu = 50$ cm/s es evidencia de que la verdadera media μ es en realidad 50 cm/s; es decir, dicha evidencia apoya la hipótesis nula H_0. Por otra parte, una media muestral que difiera considerablemente de 50 cm/s es evidencia en favor de la hipótesis alternativa H_1. Por tanto, en este caso la media muestral es el estadístico de la prueba.

La media muestral puede tomar muchos valores diferentes. Suponga que si $48.5 \leq \bar{x} \leq 51.5$, no se rechazará la hipótesis nula H_0: $\mu = 50$, y cuando $\bar{x} < 48.5$ o bien $\bar{x} > 51.5$, se recha-zará la hipótesis nula en favor de la hipótesis alternativa H_1: $\mu \neq 50$. Esto se ilustra en la figura 8-1. Los valores de $\bar{x}$ que son menores que 48.5 y mayores que 51.5 constituyen la **región crítica** de la prueba, en tanto que todos los valores que están en el intervalo $48.5 \leq \bar{x} \leq 51.5$ forman una región en la que no la hipótesis nula no puede rechazarse. Por convención, a ésta suele llamársele la **región de aceptación**. A los límites entre las regiones críticas y la región de aceptación se les llama los **valores críticos**. En el ejemplo anterior los valores críticos son 48.5 y 51.5. Se acostumbra enunciar las conclusiones en términos de la hipótesis nula H_0. Por lo tanto, se rechaza H_0 en favor de H_1 cuando el estadístico de prueba cae en la región crítica, mientras que en caso contrario no puede rechazarse H_0.

Este procedimiento de decisión puede llevar a una de dos conclusiones incorrectas. Por ejemplo, la verdadera media de la rapidez de combustión de la carga propulsora podría ser igual a 50 cm/s. Sin embargo, para las observaciones de la carga propulsora seleccionadas al azar que se prueban, pudo haberse encontrado un valor del estadístico de prueba $\bar{x}$ que caiga en la región crítica. En tal caso, se rechazaría la hipótesis nula H_0 en favor de la hipótesis alternativa H_1 cuando, de hecho, H_0 en realidad es verdadera. A este tipo de conclusión incorrecta se le llama **error tipo I**.

Definición

El **error tipo I** se define como el rechazo de la hipótesis nula H_0 cuando ésta es verdadera.

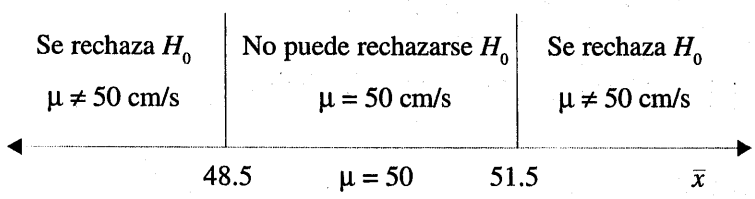


Figura 8-1 Criterios de decisión para probar H_0: $\mu = 50$ cm/s contra H_1: $\mu \neq 50$ cm/s.

Suponga ahora que la verdadera media de la velocidad de combustión es diferente de 50 cm/s, no obstante que la media muestral $\bar{x}$ cae en la región de aceptación. En este caso no se rechazaría H_0 cuando es falsa. A este tipo de conclusión incorrecta se le llama **error tipo II**.

Definición

El **error tipo II** se define como la aceptación de la hipótesis nula cuando ésta es falsa.

Por tanto, cuando se prueba cualquier hipótesis estadística, cuatro situaciones diferentes determinan si la decisión final es correcta o errónea. En la tabla 8-1 se presentan estas situaciones.

Debido a que la decisión que se tome está basada en variables aleatorias, pueden asociarse probabilidades con los errores tipo I y tipo II de la tabla 8-1. La probabilidad de incurrir en un error tipo I se denota por la letra griega α. Es decir,

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 \text{ cuando } H_0 \text{ es verdadera}) \tag{8-3}$$

En ocasiones a la probabilidad del error tipo I se le llama el **nivel de significación** o **amplitud** de la prueba. En el ejemplo de la rapidez de combustión de la carga propulsora, ocurrirá un error tipo I cuando $\bar{x} > 51.5$ o cuando $\bar{x} < 48.5$, dado que la verdadera media de la rapidez de combustión es $\mu = 50$ cm/s. Suponga que la desviación estándar de la rapidez de combustión es $\sigma = 2.5$ cm/s y que la rapidez de combustión tiene una distribución para la que se cumplen las condiciones del teorema del límite central, de tal modo que la distribución de la media muestral es aproximadamente normal con media $\mu = 50$ y desviación estándar $\sigma/\sqrt{n} = 2.5/\sqrt{10} = 0.79$. La probabilidad de incurrir en un error tipo I (es decir, el nivel de significación de la prueba) es igual a la suma de las áreas que se han sombreado en las colas de la distribución normal de la figura 8-2. Esta probabilidad puede encontrarse como

$$\alpha = P(\bar{X} < 48.5 \text{ cuando } \mu = 50) + P(\bar{X} > 51.5 \text{ cuando } \mu = 50)$$

Los valores z que corresponden a los valores críticos 48.5 y 51.5 son

$$z_1 = \frac{48.5 - 50}{0.79} = -1.90$$

Tabla 8-1 Decisiones en la prueba de hipótesis

Decisión	H_0 es verdadera	H_0 es falsa
No puede rechazarse H_0	no hay error	error tipo II
Se rechaza H_0	error tipo I	no hay error

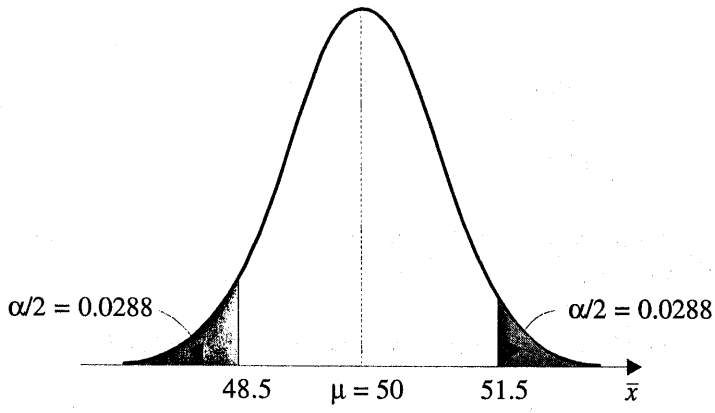


Figura 8-2 La región crítica para H_0: $\mu = 50$ contra H_1: $\mu \neq 50$ y $n = 10$.

y

$$z_2 = \frac{51.5 - 50}{0.79} = 1.90$$

Por lo tanto,

$$\begin{aligned} \alpha &= P(Z < -1.90) + P(Z > 1.90) \\ &= 0.0288 + 0.0288 \\ &= 0.0576 \end{aligned}$$

Esto implica que 5.76% de todas las muestras aleatorias llevarían al rechazo de la hipótesis H_0: $\mu = 50$ cm/s cuando la verdadera media de la rapidez de combustión es en realidad 50 cm/s.

Al examinar la figura 8-2, obsérvese que α puede reducirse ampliando la región de aceptación. Por ejemplo, si los valores críticos se hacen 48 y 52, el valor de α es

$$\begin{aligned} \alpha &= P\left(Z < \frac{48 - 50}{0.79}\right) + P\left(Z > \frac{52 - 50}{0.79}\right) \\ &= P(Z < -2.53) + P(Z > 2.53) \\ &= 0.0057 + 0.0057 \\ &= 0.0114 \end{aligned}$$

También podría reducirse α incrementando el tamaño de la muestra. Si $n = 16$, entonces $\sigma/\sqrt{n} = 2.5/\sqrt{16} = 0.625$, y al utilizar la región crítica original de la figura 8-1 se encuentra

$$z_1 = \frac{48.5 - 50}{0.625} = -2.40$$

y

$$z_2 = \frac{51.5 - 50}{0.625} = 2.40$$

Por lo tanto,

$$\begin{aligned}\alpha &= P(Z < -2.40) + P(Z > 2.40)\\ &= 0.0082 + 0.0082\\ &= 0.0164\end{aligned}$$

Al evaluar un procedimiento de prueba de hipótesis, también es importante examinar la probabilidad de un error tipo II, la cual se denotará por β. Es decir,

$$\beta = P(\text{error tipo II}) = P(\text{no puede rechazarse } H_0 \text{ cuando } H_0 \text{ es falsa}) \tag{8-4}$$

Para calcular β es necesario tener una hipótesis alternativa específica; es decir, debe tenerse un valor particular de μ. Por ejemplo, suponga que es importante rechazar la hipótesis nula H_0: $\mu = 50$ siempre que la rapidez de combustión media μ sea mayor que 52 cm/s o menor que 48 cm/s. Podría calcularse la probabilidad β del error tipo II para los valores $\mu = 52$ y $\mu = 48$, y usar este resultado para tener una idea de cuál sería el desempeño del procedimiento de prueba. Específicamente, ¿cómo funcionaría el procedimiento de prueba si se deseara detectar, es decir, rechazar H_0 para un valor de la media de $\mu = 52$ o $\mu = 48$? Debido a la simetría, sólo es necesario evaluar uno de los dos casos; por ejemplo, encontrar la probabilidad de aceptar la hipótesis nula H_0: $\mu = 50$ cm/s cuando la verdadera media es $\mu = 52$ cm/s.

La figura 8-3 ayudará a calcular la probabilidad β del error tipo II. La distribución normal de la izquierda en la figura 8-3 es la distribución del estadístico de la prueba $\bar{X}$ cuando la hipótesis nula H_0: $\mu = 50$ es verdadera (esto es lo que quiere decir la expresión "bajo H_0: $\mu = 50$"), y la distribución normal de la derecha es la distribución de $\bar{X}$ cuando la hipótesis alternativa es verdadera y el valor de la media es 52 (o "bajo H_1: $\mu = 52$"). Entonces se incurrirá en un error tipo II si la media muestral $\bar{x}$ está entre 48.5 y 51.5 (las fronteras de la región crítica) cuando

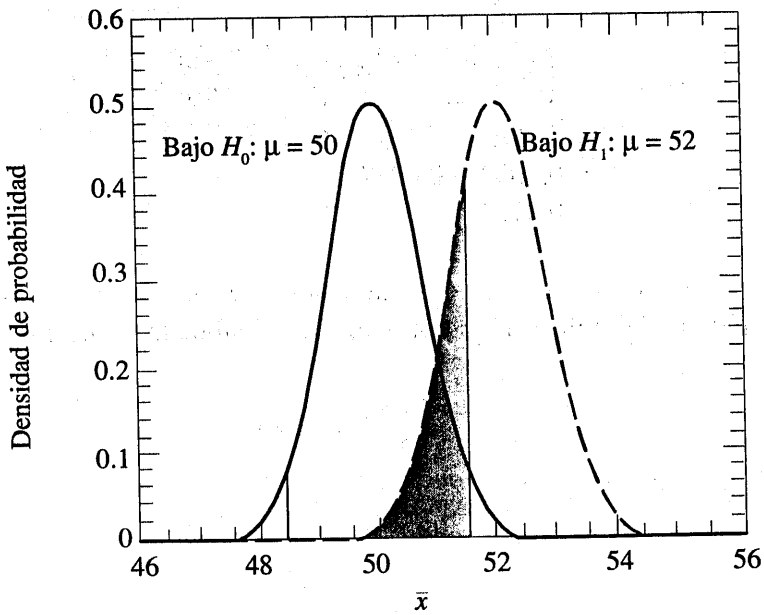


Figura 8-3 La probabilidad del error tipo II cuando $\mu = 52$ y $n = 10$.

$\mu = 52$. Como se ve en la figura 8-3, ésta es simplemente la probabilidad de que $48.5 \leq \bar{X} \leq 51.5$ cuando la verdadera media es $\mu = 52$, o el área sombreada bajo la distribución normal de la derecha. Por lo tanto, con referencia a la figura 8-3, se encuentra que

$$\beta = P(48.5 \leq \bar{X} \leq 51.5 \text{ cuando } \mu = 52)$$

Los valores z correspondientes a 48.5 y 51.5 cuando $\mu = 52$ son

$$z_1 = \frac{48.5 - 52}{0.79} = -4.43$$

y

$$z_2 = \frac{51.5 - 52}{0.79} = -0.63$$

Por lo tanto,

$$\begin{aligned} \beta &= P(-4.43 \leq Z \leq -0.63) \\ &= P(Z \leq -0.63) - P(Z \leq -4.43) \\ &= 0.2643 - 0.000 \\ &= 0.2643 \end{aligned}$$

Por tanto, si se está probando H_0: $\mu = 50$ contra H_1: $\mu \neq 50$ con $n = 10$, y el verdadero valor de la media es $\mu = 52$, la probabilidad de que no pueda rechazarse la hipótesis nula falsa es 0.2643. Por simetría, si el verdadero valor de la media es $\mu = 48$, el valor de β también será 0.2643.

La probabilidad β de incurrir en un error tipo II aumenta con rapidez conforme el verdadero valor de μ se aproxima al valor propuesto en la hipótesis. Por ejemplo, ver la figura 8-4, donde

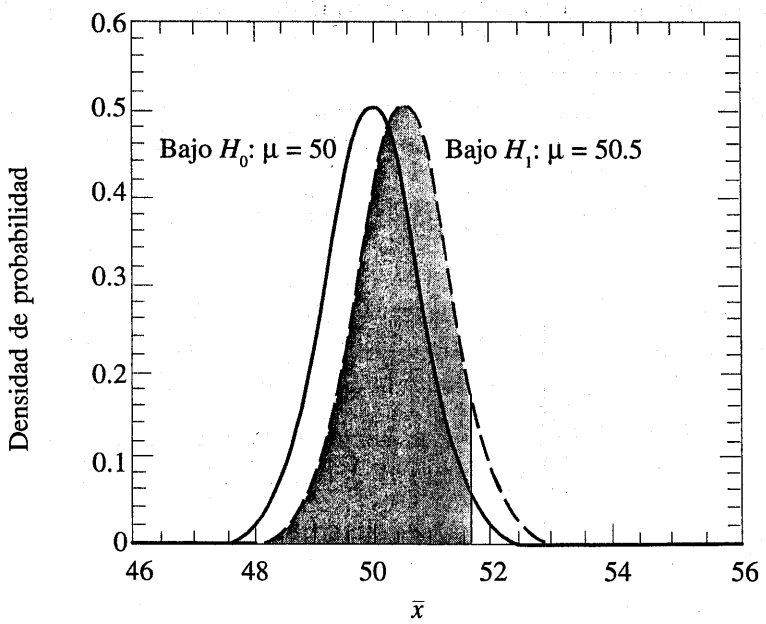


Figura 8-4 La probabilidad del error tipo II cuando $\mu = 50.5$ y $n = 10$.

el verdadero valor de la media es $\mu = 50.5$ y el valor propuesto en la hipótesis es H_0: $\mu = 50$. El verdadero valor de μ está muy cerca de 50, y el valor de β es

$$\beta = P(48.5 \leq \bar{X} \leq 51.5 \text{ cuando } \mu = 50.5)$$

Como se muestra en la figura 8-4, los valores z correspondientes a 48.5 y 51.5 cuando $\mu = 50.5$ son

$$z_1 = \frac{48.5 - 50.5}{0.79} = -2.53$$

y

$$z_2 = \frac{51.5 - 50.5}{0.79} = 1.27$$

Por lo tanto,

$$\begin{aligned}\beta &= P(-2.53 \leq Z \leq 1.27)\\ &= P(Z \leq 1.27) - P(Z \leq -2.53)\\ &= 0.8980 - 0.0057\\ &= 0.8923\end{aligned}$$

Así, la probabilidad del error tipo II es mucho más alta para el caso en que la verdadera media es 50.5 cm/s que para el caso en que la media es 52 cm/s. Desde luego, en muchas situaciones prácticas no nos preocuparía mucho incurrir en un error tipo II si la media estuviera "cerca" del valor hipotético propuesto. Nos interesaría mucho más detectar diferencias significativas entre la verdadera media y el valor especificado en la hipótesis nula.

La probabilidad del error tipo II también depende del tamaño de la muestra n. Suponga que la hipótesis nula es H_0: $\mu = 50$ cm/s y que el verdadero valor de la media es $\mu = 52$. Si el tamaño de la muestra se incrementa de $n = 10$ a $n = 16$, se tiene como resultado la situación ilustrada en la figura 8-5. La distribución normal de la izquierda es la distribución de $\bar{X}$ cuando la media es

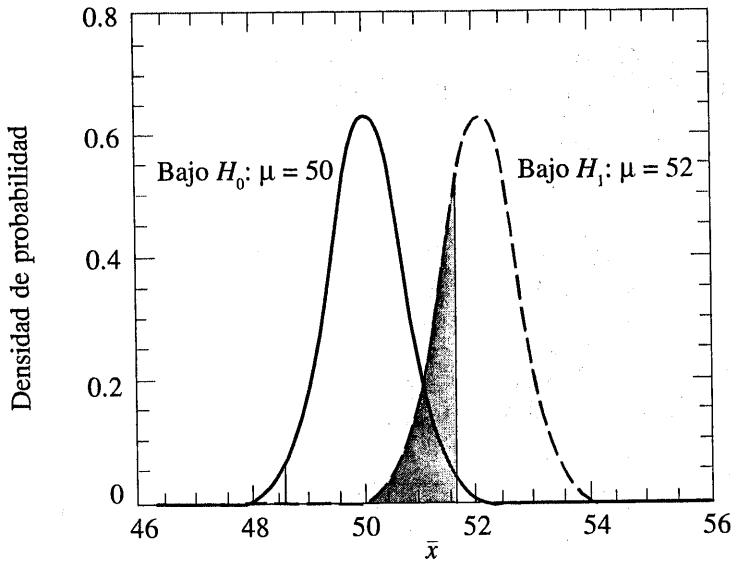


Figura 8-5 La probabilidad del error tipo II cuando $\mu = 52$ y $n = 16$.

$\mu = 50$, y la distribución normal de la derecha es la distribución de $\bar{X}$ cuando $\mu = 52$. Como se muestra en la figura 8-5, la probabilidad del error tipo II es

$$\beta = P(48.5 \leq \bar{X} \leq 51.5 \text{ cuando } \mu = 52)$$

Cuando $n = 16$, la desviación estándar de $\bar{X}$ es $\sigma/\sqrt{n} = 2.5/\sqrt{16} = 0.625$, y los valores z correspondientes a 48.5 y 51.5 cuando $\mu = 52$ son

$$z_1 = \frac{48.5 - 52}{0.625} = -5.60 \qquad \text{y} \qquad z_2 = \frac{51.5 - 52}{0.625} = -0.80$$

Por lo tanto,

$$\begin{aligned} \beta &= P(-5.60 \leq Z \leq -0.80) \\ &= P(Z \leq -0.80) - P(Z \leq -5.60) \\ &= 0.2119 - 0.000 \\ &= 0.2119 \end{aligned}$$

Recuérdese que cuando $n = 10$ y $\mu = 52$ se encontró que $\beta = 0.2643$; por lo tanto, al incrementarse el tamaño de la muestra se produce una disminución en la probabilidad del error tipo II.

Los resultados de esta sección y algunos cálculos similares más se resumen a continuación:

Región de aceptación	Tamaño de la muestra	α	β con $\mu = 52$	β con $\mu = 50.5$
$48.5 < \bar{x} < 51.5$	10	0.0576	0.2643	0.8923
$48 < \bar{x} < 52$	10	0.0114	$\boxed{0.5000}$	$\boxed{0.9705}$
$48.5 < \bar{x} < 51.5$	16	0.0164	0.2119	$\boxed{0.9445}$
$48 < \bar{x} < 52$	16	$\boxed{0.0014}$	$\boxed{0.5000}$	$\boxed{0.9918}$

Los resultados encerrados en rectángulos no se calcularon en el texto, pero el lector puede comprobarlos fácilmente. Esta tabla y la revisión anterior revelan cuatro puntos importantes:

1. El tamaño de la región crítica, y por consiguiente la probabilidad α de un error tipo I, siempre puede reducirse mediante la selección apropiada de los valores críticos.
2. Los errores tipo I y tipo II están relacionados. Una disminución de la probabilidad de un tipo de error siempre resulta en un incremento de la probabilidad del otro, siempre que no cambie el tamaño de la muestra n.
3. Un incremento del tamaño de la muestra generalmente dará como resultado una reducción tanto de α como de β, siempre que los valores críticos se mantengan constantes.
4. Cuando la hipótesis nula es falsa, β se incrementa conforme el verdadero valor del parámetro se aproxima al valor hipotético propuesto en la hipótesis nula. El valor de β disminuye conforme se incrementa la diferencia entre la verdadera media y el valor hipotético propuesto.

En general, el analista controla la probabilidad α del error tipo I cuando selecciona los valores críticos. Por tanto, suele ser sencillo que el analista establezca la probabilidad del error tipo I en (o cerca de) cualquier valor deseado. Puesto que el analista puede controlar directamente la probabilidad de rechazar incorrectamente H_0, al rechazo de la hipótesis nula H_0 siempre se le considera una **conclusión robusta**.

Por otra parte, la probabilidad β del error tipo II no es una constante, sino que depende del verdadero valor del parámetro. Depende también del tamaño de la muestra que se haya seleccionado. Debido a que la probabilidad β del error tipo II es una función tanto del tamaño de la muestra como de la medida en que la hipótesis nula H_0 es falsa, se acostumbra considerar la decisión de aceptar H_0 como una **conclusión débil**, a menos que se sepa que β es aceptablemente pequeña. Por lo tanto, en vez de decir "se acepta H_0", se prefiere la terminología "no puede rechazarse H_0". No poder rechazar H_0 implica que no se ha encontrado evidencia suficiente para rechazar H_0, es decir, para hacer una afirmación robusta. No poder rechazar H_0 no significa necesariamente que exista una alta probabilidad de que H_0 sea verdadera. Puede significar simplemente que se necesitan más datos para llegar a una conclusión robusta. Esto puede tener implicaciones importantes en la formulación de hipótesis.

Un concepto importante del que se hará uso es la **potencia** de una prueba estadística.

Definición

La **potencia** de una prueba estadística es la probabilidad de rechazar la hipótesis nula H_0 cuando la hipótesis alternativa es verdadera.

La potencia se calcula como $1 - \beta$, y puede interpretarse como *la probabilidad de rechazar correctamente una hipótesis nula falsa.* Es común comparar las pruebas estadísticas comparando sus propiedades de potencia. Por ejemplo, considérese el problema de la rapidez de combustión de la carga propulsora cuando se prueba H_0: $\mu = 50$ cm/s contra H_1: $\mu \neq 50$ cm/s. Suponga que el verdadero valor de la media es $\mu = 52$. Cuando $n = 10$, se encuentra que $\beta = 0.2643$, por lo que la potencia de esta prueba es $1 - \beta = 1 - 0.2643 = 0.7357$ cuando $\mu = 52$.

La potencia es una medida muy descriptiva y concisa de la **sensibilidad** de una prueba estadística, donde por sensibilidad se entiende la habilidad de la prueba para detectar diferencias; en este caso, la sensibilidad de la prueba para detectar la diferencia entre la rapidez de combustión media de 50 cm/s y 52 cm/s es 0.7357. Es decir, si la media verdadera es en realidad 52 cm/s, esta prueba rechazará correctamente H_0: $\mu = 50$ y "detectará" esta diferencia 73.57% de las veces. Si este valor de la potencia se considera muy bajo, el analista puede incrementar α o bien el tamaño de la muestra n.

8-1.3 Hipótesis de una y dos colas

A una prueba de cualquier hipótesis tal que

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

se le llama una prueba **de dos colas**, ya que es importante detectar las diferencias entre el valor propuesto hipotéticamente de la media μ_0, de los que están a ambos lados de μ_0. En una prueba como ésta, la región crítica se divide en dos partes, asignando (por lo general) una probabilidad igual a cada cola de la distribución del estadístico de la prueba.

Muchos problemas de prueba de hipótesis incluyen de manera natural una hipótesis alternativa **de una cola**, como por ejemplo

$$H_0: \mu = \mu_0$$
$$H_1: \mu > \mu_0 \tag{8-5}$$

o bien

$$H_0: \mu = \mu_0$$
$$H_1: \mu < \mu_0 \tag{8-6}$$

Si la hipótesis alternativa es H_1: $\mu > \mu_0$, la región crítica deberá estar en la cola superior de la distribución del estadístico de la prueba, mientras que si la hipótesis alternativa es H_1: $\mu < \mu_0$, la región crítica deberá estar en la cola inferior de la distribución. Por consiguiente, a estas pruebas en ocasiones se les llama pruebas **de una cola**. Por lo general es sencillo determinar la localización de la región crítica en las pruebas de una cola. Tan sólo se visualiza el comportamiento del estadístico de la prueba si la hipótesis nula es verdadera y la región crítica se sitúa en la cola o extremo apropiado de la distribución. En general, el signo de la desigualdad de la hipótesis alternativa "apunta" en la dirección de la región crítica.

Cuando se construyen hipótesis, la hipótesis nula siempre se enuncia como una igualdad, para que la probabilidad α del error tipo I pueda controlarse en un valor específico. La hipótesis alternativa podría ser de una o dos colas, dependiendo de la conclusión a la que quiera llegarse si H_0 es rechazada. Si el objetivo es hacer una afirmación que incluya expresiones tales como "mayor que", "menor que", "superior a", "que exceda", "al menos", etc., entonces es apropiada una hipótesis alternativa de una cola. Si la afirmación no implica ninguna dirección, o si va a hacerse una afirmación del tipo "diferente de", entonces deberá usarse una hipótesis alternativa de dos colas.

EJEMPLO 8-1

Considérese el problema de la rapidez de combustión de la carga propulsora. Suponga que si la rapidez de combustión es menor que 50 cm/s, se querría demostrar este hecho con una conclusión robusta. Las hipótesis deberán enunciarse como

$$H_0: \mu = 50 \text{ cm/s}$$
$$H_1: \mu < 50 \text{ cm/s}$$

Aquí, la región crítica está en la cola inferior de la distribución de $\overline{X}$. Puesto que el rechazo de H_0 siempre es una conclusión robusta, este postulado de las hipótesis producirá el resultado deseado si H_0 se rechaza. Obsérvese que, aun cuando la hipótesis nula está enunciada con el signo igual, se sobreentiende que incluye cualquier valor de μ no especificado en la hipótesis alternativa. Por lo tanto, el hecho de que H_0 no pueda rechazarse no quiere decir que $\mu = 50$ cm/s exactamente, sino tan sólo que no se cuenta con evidencia fuerte para sustentar H_1.

En algunos problemas del mundo real donde los procedimientos de prueba de una cola son los indicados, en ocasiones resulta difícil elegir una formulación apropiada de la hipótesis alternativa. Por ejemplo, suponga que una compañía embotelladora de refrescos compra botellas de 10 onzas a un fabricante de envases de vidrio. El embotellador quiere asegurarse de que las botellas cumplen con las especificaciones de presión interna o resistencia a reventar, que para las botellas de 10 onzas es una resistencia mínima de 200 psi. La embotelladora ha decidido formular el procedimiento de decisión para un lote específico de botellas como un problema de hipótesis. Hay dos formulaciones posibles para este problema

$$H_0: \mu = 200 \text{ psi}$$
$$H_1: \mu > 200 \text{ psi} \tag{8-7}$$

o bien

$$H_0: \mu = 200 \text{ psi}$$
$$H_1: \mu < 200 \text{ psi} \tag{8-8}$$

Consideremos la formulación de la ecuación 8-7. Si la hipótesis nula es rechazada, las botellas se juzgarán de manera satisfactoria; si H_0 no es rechazada, la implicación es que las botellas no cumplen con las especificaciones y no deberán utilizarse. Puesto que rechazar H_0 es una conclusión robusta, esta formulación obliga al fabricante de las botellas a "demostrar" que la media de la resistencia a reventar de las botellas excede la especificación. Consideremos ahora la formulación de la ecuación 8-8. En esta situación, las botellas pueden juzgarse satisfactorias a menos que se rechace H_0. Es decir, se concluye que las botellas son satisfactorias a menos que haya fuertes evidencias de lo contrario.

¿Cuál de las dos formulaciones, la de la ecuación 8-7 o la de la ecuación 8-8, es la correcta? La respuesta es, "depende". Para la ecuación 8-7, hay ciertas probabilidades de que H_0 no será rechazada (es decir, de que se decida que las botellas no son satisfactorias), aun cuando la verdadera media sea ligeramente mayor que 200 psi. Esta formulación implica que la intención es que el fabricante de botellas demuestre que el producto cumple con nuestras especificaciones o las excede. Esta formulación podría ser apropiada si el fabricante de botellas ha tenido dificultades para cumplir con las especificaciones en el pasado, o si las consideraciones de seguridad del producto nos obligan a apegarnos estrictamente a la especificación de 200 psi. Por otra parte, para la formulación de la ecuación 8-8, hay ciertas probabilidades de que H_0 sea aceptada y las botellas se juzguen satisfactorias, aun cuando la verdadera media sea ligeramente menor que 200 psi. Las botellas sólo se considerarían insatisfactorias cuando haya evidencias fuertes de que la media no excede 200 psi, es decir, cuando es rechazada H_0: $\mu = 200$ psi. Esta formulación supone que la embotelladora está relativamente satisfecha con el desempeño pasado del fabricante de botellas y que las desviaciones ligeras de la especificación $\mu \geq 200$ psi no son riesgosas.

Cuando se formulen hipótesis alternativas de una cola, deberá recordarse que rechazar H_0 siempre es una conclusión robusta. Por consiguiente, el postulado acerca de lo que es importante para llegar a una conclusión robusta deberá colocarse en la hipótesis alternativa. En los problemas del mundo real, esto dependerá por lo general del criterio del analista y de su experiencia con la situación.

8-1.4 Procedimiento general de las pruebas de hipótesis

En el presente capítulo se desarrollan los procedimientos de prueba de hipótesis para diversos problemas prácticos. Se recomienda el uso de la siguiente serie de pasos al aplicar la metodología para la prueba de hipótesis.

1. Por el contexto del problema, identificar el parámetro de interés.
2. Establecer la hipótesis nula H_0.
3. Especificar una hipótesis alternativa apropiada, H_1.
4. Elegir un nivel de significación α.
5. Establecer un estadístico de la prueba apropiado.
6. Establecer la región de rechazo del estadístico.
7. Calcular las cantidades muestrales necesarias, sustituirlas en la ecuación del estadístico de la prueba, y calcular ese valor.
8. Decidir si deberá rechazarse o no H_0 y contextualizar la decisión en el problema.

Los pasos 1 al 4 deberán completarse antes de pasar al examen de los datos muestrales. Esta serie de pasos se ilustrará en las secciones siguientes.

EJERCICIOS DE LA SECCIÓN 8-1

8-1. Un fabricante de fibras textiles investiga una nueva fibra para cortinas. La compañía sostiene que la fibra tiene una elongación media de 12 kg con una desviación estándar de 0.5 kg. La compañía desea probar la hipótesis H_0: $\mu = 12$ contra H_1: $\mu < 12$, para lo cual utiliza una muestra aleatoria de cuatro observaciones.

a) ¿Cuál es la probabilidad del error tipo I si la región crítica se define como $\bar{x} < 11.5$ kg?

b) Encontrar β para el caso en que la media verdadera de la elongación es 11.25 kg.

8-2. Repita el ejercicio 8-1 utilizando un tamaño de la muestra $n = 16$ y la misma región crítica.

8-3. En el ejercicio 8-1, encuentre el límite de la región crítica si la probabilidad del error tipo I se especifica como a = 0.01.

8-4. En el ejercicio 8-2, encuentre el límite de la región crítica si la probabilidad del error tipo I se especifica como 0.05.

8-5. El calor, expresado en calorías por gramo, que se desprende de una mezcla de cemento tiene una distribución aproximadamente normal. Se piensa que la media es 100 y que la desviación estándar es 2. Quiere probarse H_0: $\mu = 100$ contra H_1: $\mu \neq 100$ con una muestra de $n = 9$ observaciones.

a) Si la región de aceptación se define como $98.5 \leq \bar{x} \leq 101.5$, encuentre la probabilidad α del error tipo I.

b) Encuentre β para el caso en que la verdadera media del calor desprendido es 103.

c) Encuentre β para el caso en que es 105 la verdadera media del calor desprendido. Este valor de β es menor que el que se encontró en el inciso b). ¿Por qué?

8-6. Repita el ejercicio 8-5 utilizando un tamaño de la muestra $n = 5$ y la misma región de aceptación.

8-7. Una compañía de artículos de consumo está formulando un nuevo champú y está interesada en la altura de la espuma (en mm). La altura de la espuma tiene una distribución aproximadamente normal y una desviación estándar de 20 mm. La compañía quiere probar H_0: $\mu = 175$ mm contra H_1: $\mu > 175$ mm, para lo cual utiliza los resultados de $n = 10$ muestras.

a) Encuentre la probabilidad α del error tipo I si la región crítica es $\bar{x} > 185$.

b) ¿Cuál es la probabilidad del error tipo II si la verdadera media de la altura de la espuma es 195 mm?

8-8. En el ejercicio 8-7, suponga que los datos muestrales dan como resultado $\bar{x} = 190$ mm.

a) ¿A qué conclusión se llegaría?

b) ¿Qué tan "raro" es el valor muestral $\bar{x} =$ 190 mm si la verdadera media es en realidad 175 mm? Es decir, ¿cuál es la probabilidad de que se observe un promedio muestral de hasta 190 mm (o mayor) si la verdadera media de la altura de la espuma fue en realidad 175 mm?

8-9. Repita el ejercicio 8-7 suponiendo que el tamaño de la muestra es $n = 16$ y el límite de la región crítica es el mismo.

8-10. Considérese el ejercicio 8-7, y suponga que el tamaño de la muestra se incrementa a $n = 16$.

a) ¿Dónde se colocaría el límite de la región crítica si la probabilidad del error tipo I debe conservar el valor que tomó cuando $n = 10$?

b) Utilizando $n = 16$ y la nueva región crítica encontrada en el inciso a), encuentre la probabilidad β del error tipo II si la verdadera media de la altura de la espuma es 195 mm.

c) Compare el valor de β obtenido en el inciso b) con el valor del ejercicio 8-7b). ¿A qué conclusiones puede llegarse?

8-11. Un fabricante está interesado en el voltaje de salida de una fuente de poder usada en una PC. Se supone que el voltaje de salida tiene una distribución normal, con desviación estándar 0.25 V, y el fabricante quiere probar H_0: $\mu = 5$ V contra H_1: $\mu \neq 5$ V, para lo cual utiliza $n = 8$ unidades.

a) La región de aceptación es $4.85 \leq \bar{x} \leq 5.15$. Encuentre el valor de α.

b) Encuentre la potencia de la prueba para detectar una media verdadera del voltaje de salida de 5.1 V.

8-12. Resuelva de nuevo el ejercicio 8-11 cuando el tamaño de la muestra es 16 y los límites de la región de aceptación no cambian.

8-13. Considérese el ejercicio 8-11, y suponga que el fabricante quiere que la probabilidad del error tipo I de la prueba sea $\alpha = 0.05$. ¿Dónde deberá localizarse la región de aceptación?

8-14. Si se grafica la probabilidad de aceptar H_0: $\mu = \mu_0$ contra varios valores de μ y se unen los puntos con una curva suave, se obtiene la **curva de operación característica** (o **curva OC**) del procedimiento de prueba. Estas curvas tienen un uso generalizado en las aplicaciones de prueba de hipótesis en la industria para mostrar la sensibilidad y el desempeño relativo de la prueba. Cuando la verdadera media en realidad es igual a μ_0, la probabilidad de aceptar H_0 es $1 - \alpha$. Construya una curva OC para el ejercicio 8-7, utilizando los valores para la verdadera media μ de 178, 181, 184, 187, 190, 193, 196 y 199.

8-15. Convierta la curva OC del ejercicio 8-14 en una gráfica de la **función de potencia** de la prueba.

8-16. A una muestra aleatoria de 500 votantes registrados de Phoenix se les pregunta si están de acuerdo con el uso de combustibles oxigenados todo el año para reducir la contaminación ambiental. Si más de 400 votantes responden afirmativamente, se concluirá que al menos 60% de los votantes están en favor del uso de estos combustibles.

a) Encuentre la probabilidad del error tipo I, si exactamente 60% de los votantes están en favor del uso de estos combustibles.

b) ¿Cuál es la probabilidad β del error tipo II si 75% de los votantes están en favor de esta medida?

[*Sugerencia:* Utilice la aproximación normal a la binomial.]

8-17. Se piensa que la proporción de los residentes de Phoenix que están en favor de construir carreteras de cuota para complementar la red estatal de autopistas es $p = 0.3$. Si una muestra aleatoria de 10 residentes indica que 1 o menos están en favor de la propuesta se concluirá que $p < 0.3$.

a) Encuentre la probabilidad del error tipo I si la verdadera proporción es $p = 0.3$.

b) Encuentre la probabilidad de incurrir en un error tipo II con este procedimiento si $p = 0.2$.
c) ¿Cuál es la potencia de este procedimiento si la verdadera proporción es $p = 0.2$?

8-18. Se estima que la proporción de los adultos que viven en Tempe, Arizona, que tienen un título universitario es $p = 0.4$. Para probar esta hipótesis, se selecciona una muestra aleatoria de 15 adultos de Tempe. Si el número de graduados universitarios está entre 4 y 8, se aceptará la hipótesis; en caso contrario, se concluirá que $p \neq 4$.
a) Encuentre la probabilidad del error tipo I de este procedimiento, suponiendo que $p = 0.4$.
b) Encuentre la probabilidad de incurrir en un error tipo II si la verdadera proporción en realidad es $p = 0.2$.

8-2 INFERENCIA SOBRE LA MEDIA DE UNA POBLACIÓN, VARIANZA CONOCIDA

En esta sección se consideran las inferencias acerca de la media μ de una sola población cuando se conoce la varianza poblacional σ^2. Se partirá del supuesto que se ha tomado una muestra aleatoria $X_1, X_2, ..., X_n$ de la población. Con base en los temas expuestos con anterioridad, la media muestral $\bar{X}$ es un **estimador puntual insesgado** de μ. Los supuestos restantes son mínimos: ya que la inferencia se basará en la media muestral $\bar{X}$, se supondrá que la población tiene una distribución normal o bien que se cumplen las condiciones del teorema del límite central. Por tanto, la distribución de $\bar{X}$ es aproximadamente normal con media μ y varianza σ^2/n.

8-2.1 Pruebas de hipótesis sobre la media

Suponga que quieren probarse las hipótesis

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0 \tag{8-9}$$

donde μ_0 es una constante dada. Se tiene una muestra aleatoria $X_1, X_2, ..., X_n$ de la población. Puesto que $\bar{X}$ tiene una distribución normal aproximada (es decir, la **distribución de muestreo** de $\bar{X}$ es aproximadamente normal) con media μ_0 y desviación estándar $\sigma/\sqrt{n}$ si la hipótesis nula es verdadera, sería posible construir una región crítica basada en el valor calculado de la media muestral $\bar{x}$, como en la sección 8-1.2.

Suele ser más conveniente *estandarizar* la media muestral y utilizar un estadístico de prueba basado en la distribución normal estándar. Es decir, en el procedimiento de prueba de H_0: $\mu = \mu_0$ se utilice el **estadístico de prueba**

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \tag{8-10}$$

Si la hipótesis nula H_0: $\mu = \mu_0$ es verdadera, entonces $E(X) = \mu_0$, y se sigue que la distribución de Z_0 es la distribución normal estándar [denotada $N(0, 1)$]. Por consiguiente, si H_0: $\mu = \mu_0$ es verdadera, la probabilidad de que el estadístico de la prueba Z_0 esté entre $-z_{\alpha/2}$ y $z_{\alpha/2}$ es $1 - \alpha$, donde $z_{\alpha/2}$ es el punto porcentual $100\alpha/2$ de la distribución normal estándar. En la figura 8-6 se ilustran las regiones asociadas con $z_{\alpha/2}$ y $-z_{\alpha/2}$. Obsérvese que la probabilidad de que el estadístico de la prueba Z_0 esté en la región $Z_0 > z_{\alpha/2}$ o $Z_0 < -z_{\alpha/2}$ es α cuando H_0: $\mu = \mu_0$ es verdadera. Evidentemente, una muestra que produzca un valor del estadístico de prueba que esté en las colas de la distribución de Z_0 sería extraño si H_0: $\mu = \mu_0$ es verdadera; por lo tanto, se trata de un indicio de que H_0 es falsa. Así, H_0 deberá rechazarse si

$$Z_0 > z_{\alpha/2} \tag{8-11}$$

o si

$$Z_0 < -z_{\alpha/2} \tag{8-12}$$

y no podrá rechazarse H_0 si

$$-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2} \tag{8-13}$$

La ecuación 8-13 define la **región de aceptación** de H_0, y las ecuaciones 8-11 y 8-12 definen la **región crítica** o **región de rechazo.** La probabilidad del error tipo I para este procedimiento de prueba es α.

En general, es más sencillo entender la región crítica y el procedimiento de prueba cuando el estadístico de la prueba es Z_0 en vez de $\bar{X}$. Sin embargo, la misma región crítica siempre puede expresarse en términos del valor calculado de la media muestral $\bar{x}$. Un procedimiento idéntico al descrito arriba es el siguiente:

$$\text{Rechazar } H_0\text{: } \mu = \mu_0 \text{ si } \bar{x} > a \text{ o si } \bar{x} < b$$

donde

$$a = \mu_0 + z_{\alpha/2}\sigma/\sqrt{n} \tag{8-14}$$

$$b = \mu_0 - z_{\alpha/2}\sigma/\sqrt{n} \tag{8-15}$$

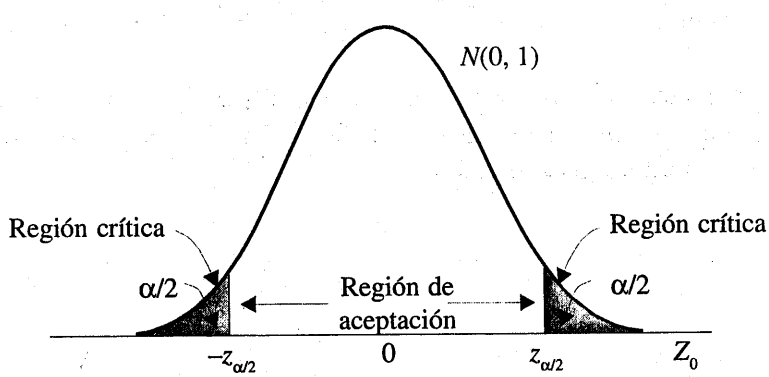


Figura 8-6 La distribución de Z_0 cuando H_0: $\mu = \mu_0$ es verdadera, con la región crítica para H_1: $\mu \neq \mu_0$.

EJEMPLO 8-2

Los sistemas de expulsión de la tripulación de un avión son impulsados por una carga propulsora sólida. La rapidez de combustión de esta carga propulsora es una característica importante del producto. Las especificaciones requieren que la rapidez de combustión media debe ser 50 cm/s. Se sabe que la desviación estándar de la rapidez de combustión es $\sigma = 2$ cm/s. El analista decide especificar una probabilidad del error tipo I o nivel de significación de $\alpha = 0.05$. Selecciona una muestra aleatoria de $n = 25$ y obtiene un promedio muestral de la rapidez de combustión de $\bar{x} = 51.3$ cm/s. ¿A qué conclusiones deberá llegarse?

Este problema puede resolverse aplicando el procedimiento de ocho pasos descrito en la sección 8-1.4, lo que resulta en

1. El parámetro de interés es μ, la media de la rapidez de combustión.
2. H_0: $\mu = 50$ cm/s
3. H_1: $\mu \neq 50$ cm/s
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

6. Se rechaza H_0 si $Z_0 > 1.96$ o si $Z_0 < -1.96$. Obsérvese que esto se deriva del paso 4, donde se especificó $\alpha = 0.05$ y, por tanto, los límites de la región crítica están en $z_{0.025} = 1.96$ y $-z_{0.025} = -1.96$.
7. Cálculos: Puesto que $\bar{x} = 51.3$ y $\sigma = 2$,

$$Z_0 = \frac{51.3 - 50}{2/\sqrt{25}} = 3.25$$

8. Conclusión: Puesto que $Z_0 = 3.25 > 1.96$, se rechaza H_0: $\mu = 50$ en el nivel de significación 0.05. Expresado con mayor detalle, se concluye que la media de la rapidez de combustión difiere de 50 cm/s, con base en una muestra de 25 mediciones. De hecho, hay evidencias fuertes de que la rapidez de combustión media excede 50 cm/s.

También pueden desarrollarse procedimientos para probar hipótesis sobre la media μ cuando la hipótesis alternativa es de una cola. Suponga que las hipótesis se especifican como

$$\begin{aligned} H_0&: \mu = \mu_0 \\ H_1&: \mu > \mu_0 \end{aligned} \tag{8-16}$$

Cuando se define la región crítica para esta prueba, se observa que un valor negativo del estadístico de la prueba Z_0 nunca nos llevaría a concluir que H_0: $\mu = \mu_0$ es falsa. Por lo tanto, la región

crítica se colocaría en la cola superior de la distribución normal estándar y H_0 se rechazaría si el valor calculado de z_0 fuera demasiado grande. Es decir, H_0 se rechazaría si

$$Z_0 > z_\alpha \tag{8-17}$$

De manera similar, para probar

$$\begin{aligned} H_0&: \mu = \mu_0 \\ H_1&: \mu < \mu_0 \end{aligned} \tag{8-18}$$

se calcularía el estadístico de la prueba Z_0, y se rechazaría H_0 si el valor de Z_0 es demasiado pequeño. Es decir, la región crítica está en la cola inferior de la distribución normal estándar, y H_0 se rechazaría si

$$Z_0 < -z_\alpha \tag{8-19}$$

8-2.2 Valores *P* en las pruebas de hipótesis

Una forma de reportar los resultados de una prueba de hipótesis es enunciando que la hipótesis nula fue rechazada o no en un valor α o nivel de significación especificado. Por ejemplo, en el problema de carga propulsora anterior, puede decirse que H_0: $\mu = 50$ se rechazó en el nivel de significación 0.05. Este enunciado de las conclusiones con frecuencia es inadecuado, pues no le da al responsable de tomar decisiones ninguna idea acerca de si el valor calculado del estadístico de prueba apenas llegó a la región de rechazo o si se adentró bastante en ella. Además, al enunciarse los resultados de esta manera se impone el nivel de significación predefinido a los demás usuarios de la información. Este enfoque puede ser insatisfactorio, pues algunos responsables de tomar decisiones podrían sentirse incómodos con los riesgos que implica $\alpha = 0.05$.

A fin de evitar estas dificultades, en la práctica se ha adoptado de manera generalizada el **enfoque del valor *P***. El valor *P* es la probabilidad de que el estadístico de la prueba tome un valor que es al menos tan extremo como el valor observado del estadístico cuando la hipótesis nula H_0 es verdadera. Por tanto, un valor *P* comunica bastante información acerca del peso de la evidencia en contra de H_0, y así el responsable de tomar decisiones puede sacar una conclusión con *cualquier* nivel de significación especificado. Se presenta a continuación la definición formal del valor *P*.

Definición

El **valor *P*** es el nivel de significación más bajo que llevaría al rechazo de la hipótesis nula H_0 con los datos dados.

Se acostumbra llamar "significativo" al estadístico de la prueba (y a los datos) cuando la hipótesis nula H_0 es rechazada; por lo tanto, el valor *P* puede considerarse como el nivel α más

bajo en el que los datos son significativos. Cuando se conoce el valor P, el responsable de las decisiones puede determinar por sí mismo qué tan significativos son los datos sin que el analista le imponga un nivel de significación preseleccionado.

Para las pruebas de la distribución normal hasta aquí tratadas es relativamente sencillo calcular el valor P. Si z_0 es el valor calculado del estadístico de la prueba, entonces el valor P es

$$P = \begin{cases} 2[1 - \Phi(|z_0|)] & \text{para una prueba de dos colas: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu \neq \mu_0 \\ 1 - \Phi(z_0) & \text{para una prueba de cola superior: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu > \mu_0 \\ \Phi(z_0) & \text{para una prueba de cola inferior: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu < \mu_0 \end{cases} \tag{8-20}$$

Aquí, $\Phi(z)$ es la función de distribución normal estándar acumulada definida en el capítulo 5. Recuérdese que $\Phi(z) = P(Z \leq z)$, donde Z es $N(0, 1)$. Para ilustrar esto, considérese el problema de la carga del propulsor del ejemplo 8-2. El valor calculado del estadístico de la prueba es $Z_0 = 3.25$, y como la hipótesis alternativa es de dos colas, el valor P es

$$\text{Valor } P = 2[1 - \Phi(3.25)] = 0.0012$$

Por tanto, H_0: $\mu = 50$ se rechazaría en cualquier nivel de significación $\alpha \geq$ valor $P = 0.0012$. Por ejemplo, H_0 se rechazaría si $\alpha = 0.01$, pero no se rechazaría si $\alpha = 0.001$.

No siempre es sencillo calcular el valor P exacto de una prueba. Sin embargo, la mayoría de los programas de computadora modernos para análisis estadísticos reportan valores P, e incluso pueden obtenerse en algunas calculadoras de mano. Se indicará también cómo aproximar el valor P. Por último, cuando se usa el enfoque del valor P, puede modificarse el paso 6 del procedimiento de prueba de hipótesis. Específicamente, no es necesario enunciar explícitamente la región crítica.

8-2.3 Error tipo II y elección del tamaño de la muestra

Al probar hipótesis, el analista selecciona directamente la probabilidad del error tipo I. Sin embargo, la probabilidad β del error tipo II depende de la elección del tamaño de la muestra. En esta sección se indicará cómo calcular la probabilidad β del error tipo II. Se explicará asimismo cómo seleccionar el tamaño de la muestra para obtener un valor especificado de β.

Cómo encontrar la probabilidad β del error tipo II

Considérese la hipótesis de dos colas:

$$H_0\text{: } \mu = \mu_0$$
$$H_1\text{: } \mu \neq \mu_0$$

Suponga que la hipótesis nula es falsa y que el verdadero valor de la media es $\mu = \mu_0 + \delta$, por ejemplo, donde $\delta > 0$. El estadístico de la prueba Z_0 es

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

$$= \frac{\overline{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}} + \frac{\delta\sqrt{n}}{\sigma}$$

Por lo tanto, la distribución de Z_0 cuando H_1 es verdadera es

$$Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right) \tag{8-21}$$

En la figura 8-7 se ilustra la distribución del estadístico de la prueba Z_0 bajo la hipótesis nula H_0 y bajo la hipótesis alternativa H_1. Al examinar esta figura se observa que si H_1 es verdadera, sólo se incurrirá en un error tipo II si $-z_{\alpha/2} \le Z_0 \le z_{\alpha/2}$, donde $Z_0 \sim N(\delta\sqrt{n}/\sigma, 1)$. Es decir, la probabilidad β del error tipo II es la probabilidad de que Z_0 esté entre $-z_{\alpha/2}$ y $z_{\alpha/2}$, *dado que H_1 es verdadera.* Esta probabilidad se muestra en la porción sombreada de la figura 8-7. Expresada matemáticamente, esta probabilidad es

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \tag{8-22}$$

donde $\Phi(z)$ denota la probabilidad a la izquierda de z en la distribución normal estándar. Obsérvese que la ecuación 8-22 se obtuvo evaluando la probabilidad de que Z_0 esté en el intervalo $[-z_{\alpha/2}, z_{\alpha/2}]$ cuando H_1 es verdadera. Por otra parte, obsérvese que la ecuación 8-22 también es válida si $\delta < 0$, debido a la simetría de la distribución normal. También es posible derivar una expresión similar a la ecuación 8-22 para una hipótesis alternativa de una cola.

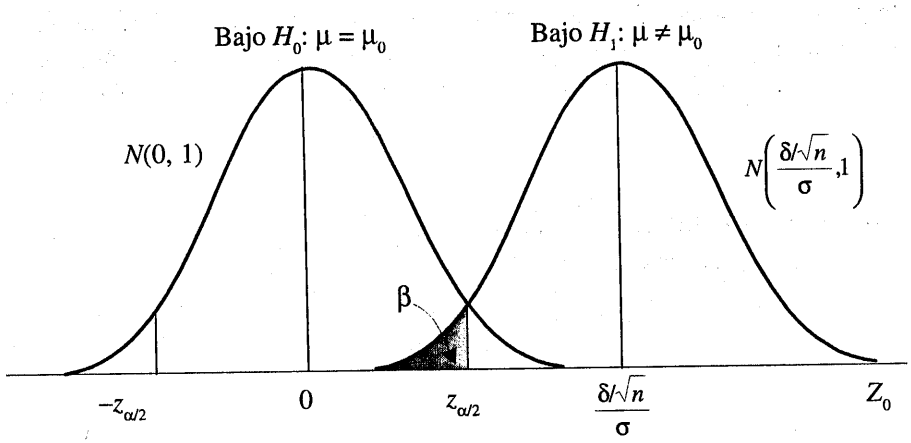


Figura 8-7 La distribución de Z_0 bajo H_0 y H_1.

Fórmulas para el tamaño de la muestra

Es sencillo obtener fórmulas que determinen el tamaño de la muestra apropiado para obtener un valor particular β para δ y α dadas. Para la hipótesis alternativa de dos colas, por la ecuación 8-22 se sabe que

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

o si $\delta > 0$,

$$\beta \simeq \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \tag{8-23}$$

puesto que $\Phi\left(-z_{\alpha/2} - \delta\sqrt{n}/\sigma\right) \simeq 0$ cuando δ es positiva. Sea z_β el percentil 100β superior de la distribución normal estándar. Entonces, $\beta = \Phi(-z_\beta)$. Por la ecuación 8-23

$$-Z_\beta \simeq Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}$$

o bien

$$n \simeq \frac{(z_{\alpha/2} + z_\beta)^2\sigma^2}{\delta^2} \tag{8-24}$$

donde

$$\delta = \mu - \mu_0$$

Esta aproximación es buena cuando $\Phi\left(-z_{\alpha/2} - \delta\sqrt{n}/\sigma\right)$ es pequeña en comparación con β. Para cualquiera de las hipótesis alternativas de una cola, el tamaño de la muestra que se requiere para producir un error tipo II especificado con probabilidad β, dadas δ y α es

$$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{\delta^2} \tag{8-25}$$

donde

$$\delta = \mu - \mu_0$$

EJEMPLO 8-3

Considérese el problema de la carga propulsora del cohete del ejemplo 8-2. Suponga que el analista desea diseñar la prueba de tal modo que si la verdadera media de la rapidez de combustión difiere de 50 cm/s a lo más por 1 cm/s, la prueba detecte esta divergencia (es decir, se rechaza H_0: $\mu = 50$) con una alta probabilidad, digamos, 0.90. Ahora bien, se observa que $\sigma = 2$, $\delta = 51 - 50 = 1$, $\alpha = 0.05$ y $\beta = 0.10$. Puesto que $z_{\alpha/2} = z_{0.025} = 1.96$ y $z_\beta = z_{0.10} = 1.28$, el tamaño de la muestra que se necesita para detectar esta divergencia de H_0: $\mu = 50$ se encuentra con la ecuación 8-24 como

$$n \simeq \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \frac{(1.96+1.28)^2 2^2}{(1)^2} \simeq 42$$

La aproximación es buena en este caso, ya que $\Phi\left(-z_{\alpha/2} - \delta\sqrt{n}/\sigma\right) = \Phi\left(-1.96 - (1)\sqrt{42}/2\right) = \Phi(-5.20) \simeq 0$, valor que es pequeño en comparación con β.

Uso de curvas de operación característica

Cuando se hacen cálculos del tamaño de la muestra o del error tipo II, en ocasiones es más conveniente utilizar las **curvas de operación característica** de las gráficas del apéndice VI*a* y VI*b*. En estas curvas se grafica β según se calcula con la ecuación 8-22 contra un parámetro d para varios tamaños de la muestra n. Se incluyen curvas tanto para $\alpha = 0.05$ como para $\alpha = 0.01$. El parámetro d se define como

$$d = \frac{|\mu - \mu_0|}{\sigma} = \frac{|\delta|}{\sigma} \tag{8-26}$$

El parámetro d se ha escogido de tal modo que pueda usarse un solo juego de curvas de operación característica para todos los problemas, independientemente de los valores de μ_0 y σ. Al examinar las curvas de operación característica, o la ecuación 8-22 y la figura 8-7, se observa que:

1. Entre más apartado esté el verdadero valor de la media μ de μ_0, menor será la probabilidad β del error tipo II para n y α dadas. Es decir, se observa que para un tamaño de la muestra y α especificados, es más fácil detectar las diferencias considerables en la media que las pequeñas.
2. Para δ y α dadas, la probabilidad β del error tipo II disminuye cuando n se incrementa. Es decir, para detectar una diferencia δ especificada en la media, la prueba puede hacerse más poderosa incrementando el tamaño de la muestra.

EJEMPLO 8-4

Considérese el problema de la carga propulsora del ejemplo 8-2. Suponga que el analista se interesa en la probabilidad del error tipo II si la verdadera media de la rapidez de combustión es $\mu = 51$ cm/s. Es posible usar las curvas de operación característica para encontrar β. Obsérvese que $\delta = 51 - 50 = 1$, $n = 25$, $\sigma = 2$ y $\alpha = 0.05$. Entonces, al aplicar la ecuación 8-26, se obtiene

$$d = \frac{|\mu - \mu_0|}{\sigma} = \frac{|\delta|}{\sigma} = \frac{1}{2}$$

y por la gráfica VI*a* del apéndice, con $n = 25$, se encuentra que $\beta = 0.30$. Es decir, si la verdadera media de la rapidez de combustión es $\mu = 51$ cm/s, entonces hay una posibilidad aproximada de 30% de que esto no se detecte en la prueba con $n = 25$.

EJEMPLO 8-5

Una vez más, considérese el problema de la carga propulsora del ejemplo 8-2. Suponga que al analista le gustaría diseñar la prueba de tal modo que si la verdadera media de la rapidez de combustión difiere de 50 cm/s a lo más por 1 cm/s, la prueba detecta esta divergencia (es decir, se rechaza H_0: $\mu = 50$) con una alta probabilidad, digamos, 0.90. Se trata exactamente del mismo requerimiento del ejemplo 8-3, donde se usó la ecuación 8-24 para encontrar que el tamaño de la muestra requerido es $n = 42$. También pueden utilizarse las curvas de operación característica para encontrar el tamaño de la muestra para esta prueba. Puesto que $d = |\mu - \mu_0|/\sigma = 1/2$, $\alpha = 0.05$ y $\beta = 0.10$, en la gráfica VI*a* del apéndice se encuentra que el tamaño de la muestra requerido es aproximadamente $n = 40$. Este resultado está cerca del tamaño de la muestra calculado con la ecuación 8-24.

En general, las curvas de operación característica incluyen tres parámetros: β, d y n. Dados cualesquiera dos de estos parámetros, es posible determinar el valor del tercero. Hay dos aplicaciones típicas de estas curvas:

1. Para n y d dados, encontrar β (como se ilustra en el ejemplo 8-3). Este tipo de problema se encuentra con frecuencia cuando al analista le preocupa la sensibilidad de un experimento que ya se realizó, o cuando el tamaño de la muestra está restringido por factores económicos o de otra índole.
2. Para β y d dados, encontrar n. Este caso se ilustró en el ejemplo 8-4. Es común encontrar este tipo de problema cuando el analista tiene la oportunidad de seleccionar el tamaño de la muestra al iniciarse un experimento.

En las gráficas VI*c* y VI*d* del apéndice se presentan las curvas de operación característica para hipótesis alternativas de una cola. Si la hipótesis alternativa es H_1: $\mu > \mu_0$, entonces la escala de las abscisas de estas cartas es

$$d = \frac{\mu - \mu_0}{\sigma} \tag{8-27}$$

Cuando la hipótesis alternativa es H_1: $\mu < \mu_0$, la escala de las abscisas correspondiente es

$$d = \frac{\mu_0 - \mu}{\sigma} \tag{8-28}$$

8-2.4 Prueba para muestras grandes

Aun cuando el procedimiento de prueba de la hipótesis nula H_0: $\mu = \mu_0$ se ha desarrollado suponiendo que σ^2 es conocida, en muchas —si no es que en la mayoría— de las situaciones prácticas no se conocerá σ^2. En general, si $n \geq 30$, la varianza muestral s^2 estará cerca de σ^2 en la mayoría de las muestras y, por tanto, s puede sustituirse por σ en los procedimientos de prueba sin efectos nocivos considerables. Por tanto, aun cuando se presentó una prueba para σ^2 conocida, ésta puede convertirse fácilmente en un *procedimiento de prueba para muestras grandes para* σ^2 *desconocida.* El tratamiento detallado del caso en que σ^2 es desconocida y n es pequeña implica el uso de la distribución t y se pospondrá hasta la sección 8-3.

8-2.5 Algunos comentarios prácticos sobre las pruebas de hipótesis

El procedimiento de ocho pasos

En la sección 8-1.4 se describió un procedimiento de ocho pasos para la prueba de hipótesis estadísticas. Este procedimiento se ilustró en el ejemplo 8-2 y volverá a encontrarse muchas veces tanto en este capítulo como en el capítulo 9. En la práctica, este procedimiento tan formal y (aparentemente) rígido no siempre es necesario. En general, cuando el experimentador (o el responsable de las decisiones) ha establecido el punto de interés y ha determinado el *diseño del experimento* (es decir, cómo van a recabarse los datos, cómo van a realizarse las mediciones y cuántas observaciones se requieren), en realidad sólo se requieren tres pasos:

1. Especificar el estadístico de la prueba que va a usarse (por ejemplo z_0).
2. Especificar la localización de la región crítica (de dos colas, de cola superior o de cola inferior).
3. Especificar los criterios de rechazo (de manera típica, el valor de α o el valor P en el que deberá ocurrir el rechazo).

Estos pasos suelen completarse de manera casi simultánea al resolver problemas del mundo real, aunque se hace hincapié en que es importante pensar con atención cada uno de los pasos. Por eso presentamos y usamos el proceso de ocho pasos: parece reforzar los aspectos fundamentales del enfoque correcto. Aun cuando no se utilice siempre al resolver problemas reales, constituye un marco de referencia útil cuando uno se inicia en el estudio de la prueba de hipótesis.

Significación estadística y significación práctica

Se señaló antes que reportar los resultados de una prueba de hipótesis en términos de un valor P es muy conveniente, pues comunica información que el enunciado simple "se rechaza H_0" o "no se puede rechazar H_0". Es decir, el rechazo de H_0 en el nivel de significación 0.05 tiene mayor sentido si el valor del estadístico de la prueba se adentra bastante en la región crítica, excediendo sobradamente el valor crítico de 5%, que si apenas excede dicho valor.

Incluso un valor P muy pequeño puede ser difícil de interpretar desde un punto de vista práctico cuando se toman decisiones, ya que, si bien un valor P pequeño indica **significación estadística** en el sentido de que deberá rechazarse H_0 en favor de H_1, la desviación real de H_0 que se ha detectado puede tener escasa (si no es que nula) **significación práctica** (los ingenieros

acostumbran decir "significación de ingeniería"). Esto es particularmente cierto cuando el tamaño de la muestra n es grande.

Por ejemplo, considérese el problema de la rapidez de combustión de la carga propulsora del ejemplo 8-3, donde se prueba H_0: $\mu = 50$ cm/s contra H_1: $\mu \neq 50$ cm/s con $\sigma = 2$. Si se supone que la rapidez de combustión media es en realidad 50.5 cm/s, entonces ésta no es una desviación seria de H_0: $\mu = 50$ cm/s en el sentido de que si la media es en realidad 50.5 cm/s, no hay ningún efecto práctico observable en el desempeño del sistema de expulsión de la tripulación. En otras palabras, concluir que $\mu = 50$ cm/s cuando es en realidad 50.5 cm/s es un error poco costoso que no tiene significación práctica. Para un tamaño de la muestra razonablemente grande, un valor real de $\mu = 50.5$ llevará a una $\bar{x}$ muestral que está cerca de 50.5 cm/s, y no se querría que este valor de $\bar{x}$ de la muestra resultara en el rechazo de H_0. A continuación se presenta el valor P para probar H_0: $\mu = 50$ cuando se observa $\bar{x} = 50.5$ cm/s y la potencia de la prueba con $\alpha = 0.05$ en el caso en que la verdadera media es 50.5 para varios tamaños de la muestra n.

Tamaño de la muestra n	Valor P cuando $\bar{x} = 50.5$	Potencia (con $\alpha = 0.05$) cuando $\mu = 50.5$
10	0.4295	0.1241
25	0.2113	0.2396
50	0.0767	0.4239
100	0.0124	0.7054
400	5.73×10^{-7}	0.9988
1 000	2.57×10^{-15}	1.0000

La columna del valor P de esta tabla indica que para tamaños grandes de la muestra, el valor muestral observado de $\bar{x} = 50.5$ indicaría claramente que deberá rechazarse H_0: $\mu = 50$, aun cuando los resultados muestrales observados en realidad no difieren mucho del valor propuesto $\mu_0 = 50$. La columna de la potencia indica que si se prueba una hipótesis en un nivel de significación fijo α y aun cuando haya poca diferencia práctica entre la verdadera media y el valor propuesto, un tamaño grande de la muestra llevará casi siempre al rechazo de H_0. La moraleja de esta demostración es clara: **es necesario ser cuidadoso cuando se interpretan los resultados de la prueba de hipótesis cuando el tamaño de la muestra es grande, ya que probablemente se detectará cualquier desviación pequeña del valor propuesto μ_0, aun cuando la diferencia tenga escasa o nula significación práctica.**

8-2.6 Intervalo de confianza para la media

En muchos casos, una estimación puntual no proporciona información suficiente acerca de un parámetro. Por ejemplo, en el problema de la carga propulsora del cohete se rechazó la hipótesis nula H_0: $\mu = 50$, y la estimación puntual de la rapidez de combustión media es $\bar{x} = 51.3$ cm/s. Sin embargo, el ingeniero preferiría tener un **intervalo** en el cual se esperaría encontrar la verdadera media de la rapidez de combustión, ya que no es factible que $\mu = 51.3$. Una forma de concluir

esto es con la estimación de un intervalo llamado **intervalo de confianza**. En la sección 7-7 se definió el concepto básico de intervalo de confianza. Recuérdese que a un intervalo tal que

$$l \leq \mu \leq u \tag{8-29}$$

se le llama intervalo de confianza del $100(1 - \alpha)$ por ciento para el parámetro μ si

$$P(L \leq \mu \leq U) = 1 - \alpha$$

Al intervalo de confianza de la ecuación 8-29 es más correcto llamarlo **intervalo de confianza de dos colas**, ya que especifica tanto un límite inferior como uno superior para μ. Ocasionalmente, podría ser más apropiado un **intervalo de confianza de una cola**. Un intervalo de confianza inferior de una cola del $100(1 - \alpha)$ por ciento para μ está dado por el intervalo

$$l \leq \mu \tag{8-30}$$

donde el límite de confianza inferior se elige de tal modo que

$$P(L \leq \mu) = 1 - \alpha \tag{8-31}$$

De manera similar, un intervalo de confianza de cola superior del $100(1 - \alpha)$ por ciento para μ está dado por el intervalo

$$\mu = u \tag{8-32}$$

donde el límite de confianza superior u se elige de tal modo que

$$P(\mu \leq U) = 1 - \alpha \tag{8-33}$$

Es muy sencillo encontrar las cantidades L y U que definen el intervalo de confianza para μ. Se sabe que la distribución de muestreo de $\overline{X}$ es normal con media μ y varianza σ^2/n. Por lo tanto, la distribución del estadístico

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

es una distribución normal estándar.

En la figura 8-8 se muestra la distribución de $Z = (\overline{X} - \mu)/(\sigma/\sqrt{n})$. Al examinar esta figura se observa que

$$P\{-z_{\alpha/2} \leq Z \leq z_{\alpha/2}\} = 1 - \alpha$$

por lo que

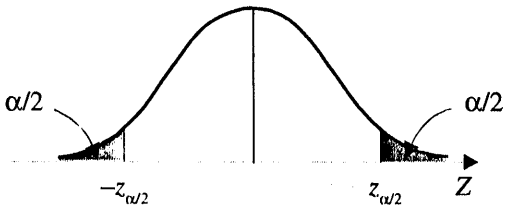


Figura 8-8 La distribución de Z.

$$P\left\{-z_{\alpha/2} \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right\} = 1 - \alpha$$

Esta expresión puede reordenarse como

$$P\left\{\overline{X} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \overline{X} + z_{\alpha/2}\sigma/\sqrt{n}\right\} = 1 - \alpha \tag{8-34}$$

Al considerar la ecuación 8-28, los límites inferior y superior de las desigualdades de la ecuación 8-34 son los límites de confianza inferior y superior L y U, respectivamente. Esto lleva a la siguiente definición:

Definición: Intervalo de confianza para la media, varianza conocida

Si $\bar{x}$ es la media muestral de una muestra aleatoria de tamaño n de una población con varianza conocida σ^2, un intervalo de confianza del $100(1 - \alpha)$ por ciento para μ está dado por

$$\bar{x} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \bar{x} + z_{\alpha/2}\sigma/\sqrt{n} \tag{8-35}$$

donde $z_{\alpha/2}$ es el punto porcentual $100\alpha/2$ superior de la distribución normal estándar.

Para muestras de una población normal, o para muestras de tamaño $n \geq 30$, independientemente de la forma de la población, el intervalo de confianza de la ecuación 8-35 dará buenos resultados. Sin embargo, para muestras pequeñas de poblaciones no normales no puede esperarse que el nivel de confianza $1 - \alpha$ sea exacto.

EJEMPLO 8-6

Considérese el problema de la carga propulsora del cohete del ejemplo 8-2. Suponga que quiere encontrarse un intervalo de confianza de 95% para la rapidez de combustión media. Puede usarse la ecuación 8-35 para construir el intervalo de confianza. Un intervalo de 95% implica que $1 - \alpha = 0.95$, por lo que $\alpha = 0.05$ y, por la tabla II del apéndice, $z_{\alpha/2} = z_{0.05/2} = z_{0.025} = 1.96$.

El límite de confianza inferior es

$$\begin{aligned} l &= x - z_{\alpha/2}\sigma/\sqrt{n} \\ &= 51.3 - 1.96(2)/\sqrt{25} \\ &= 51.3 - 0.78 \\ &= 50.52 \end{aligned}$$

y el límite de confianza superior es

$$\begin{aligned} l &= \bar{x} + z_{\alpha/2}\sigma/\sqrt{n} \\ &= 51.3 - 1.96(2)/\sqrt{25} \\ &= 51.3 - 0.78 \\ &= 50.52 \end{aligned}$$

Por tanto, el intervalo de confianza de dos colas de 95% es

$$50.52 \leq \mu \leq 52.08$$

Este es el intervalo de valores razonables para la rapidez de combustión media con una confianza de 95%.

Relación entre pruebas de hipótesis e intervalos de confianza

Existe una relación cercana entre la prueba de una hipótesis acerca de cualquier parámetro, digamos θ, y el intervalo de confianza para θ. Si $[l, u]$ es un intervalo de confianza del $100(1 - \alpha)$ por ciento para el parámetro θ, entonces la prueba de tamaño α de la hipótesis

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta \neq \theta_0 \end{aligned}$$

llevará al rechazo de H_0 si y sólo si θ_0 no está en el intervalo de confianza del $100(1 - \alpha)$ por ciento $[l, u]$. Para ejemplificar, considérese el problema de la carga propulsora del sistema de expulsión que se ha venido discutiendo. La hipótesis nula H_0: $\mu = 50$ fue rechazada, utilizando $\alpha = 0.05$. El intervalo de confianza de dos colas de 95% para μ es $50.52 \leq \mu \leq 52.08$. Es decir, el intervalo $[l, u]$ es [50.52, 52.08], y como $\mu_0 = 50$ no está incluida en este intervalo, la hipótesis nula H_0: $\mu = 50$ se rechaza.

Nivel de confianza y precisión de la estimación

Obsérvese en el ejemplo anterior que la elección del nivel de confianza de 95% fue esencialmente arbitraria. ¿Qué habría ocurrido si se hubiera elegido un nivel de confianza más alto, digamos, 99%? De hecho, ¿no parece razonable que se quisiera tener el nivel de confianza más alto? Con $\alpha = 0.01$, se encuentra $z_{\alpha/2} = z_{0.01/2} = z_{0.005} = 2.58$, en tanto que para $\alpha = 0.05$, $z_{0.025} = 1.96$. Por tanto, la longitud del intervalo de confianza del 95% es

$$2\left(1.96\sigma/\sqrt{n}\right) = 3.92\sigma/\sqrt{n}$$

mientras que la longitud del intervalo de confianza de 99% es

$$2\left(2.58\sigma/\sqrt{n}\right) = 5.16\sigma/\sqrt{n}$$

El intervalo de confianza de 99% es más largo que el intervalo de confianza de 95%. A esto se debe que el nivel de confianza en el intervalo de confianza de 99% sea más alto. En general, para un tamaño de la muestra n y la desviación estándar σ fijos, entre más alto sea el nivel de confianza, mayor será la longitud del intervalo de confianza resultante.

Puesto que la mitad de la longitud del intervalo de confianza mide la precisión de la estimación, se observa que la precisión tiene una relación inversa con el nivel de confianza. Como ya se señaló, es deseable obtener un intervalo de confianza que sea lo suficientemente corto para fines de toma de decisiones pero que ofrezca a la vez una confianza adecuada. Una forma de conseguir esto es eligiendo el tamaño de la muestra n lo suficientemente larga para obtener un intervalo de confianza de longitud especificada con confianza preestablecida.

Elección del tamaño de la muestra

La precisión del intervalo de confianza de la ecuación 8-35 es $z_{\alpha/2}\sigma/\sqrt{n}$. Esto significa que al usar $\bar{x}$ para estimar μ, el error $E = |\bar{x} - \mu|$ es menor o igual que $z_{\alpha/2}\sigma/\sqrt{n}$ con una confianza de $100(1 - \alpha)$. En la figura 8-9 se muestra esto gráficamente. En situaciones en las que el tamaño de la muestra puede controlarse, n puede escogerse de tal modo que se tenga una confianza del $100(1 - \alpha)$ por ciento de que el error al estimar μ es menor que un error especificado E. El tamaño de la muestra apropiado se encuentra escogiendo n tal que $z_{\alpha/2}\sigma/\sqrt{n} = E$. Al resolverse esta ecuación se obtiene la siguiente fórmula para n.

Definición

Si se usa $\bar{x}$ como estimación de μ, puede tenerse una confianza del $100(1 - \alpha)$ por ciento de que el error $|\bar{x} - \mu|$ no excederá una cantidad especificada E cuando el tamaño de la muestra es

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 \tag{8-36}$$

Si el miembro derecho de la ecuación 8-36 no es un entero, deberá redondearse. Con esto se asegurará que el nivel de confianza no esté abajo del $100(1 - \alpha)$ por ciento. Obsérvese que $2E$ es la longitud del intervalo de confianza resultante.

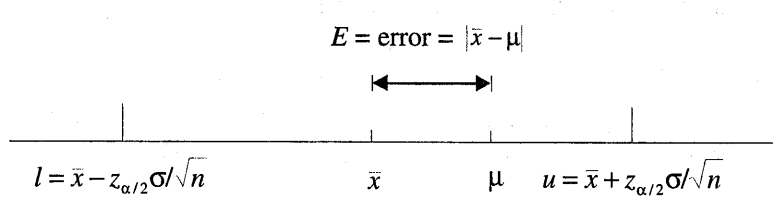


Figura 8-9 Error al estimar μ con $\bar{x}$.

EJEMPLO 8-7

Para ilustrar el uso de este procedimiento, suponga que se quisiera que el error al estimar la rapidez de combustión media de la carga propulsora del cohete fuera menor que 1.5 cm/s, con una confianza de 95%. Puesto que $\sigma = 2$ y $z_{0.025} = 1.96$, el tamaño de la muestra requerido puede encontrarse con la ecuación 8-36 como

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 = \left[\frac{(1.96)2}{1.5}\right]^2 = 6.83 \simeq 7$$

Obsérvese la relación general entre el tamaño de la muestra, la longitud deseada del intervalo de confianza $2E$, el nivel de confianza del $100(1-\alpha)$ por ciento, y la desviación estándar σ:

- Cuando la longitud deseada del intervalo $2E$ disminuye, el tamaño de la muestra n requerido se incrementa para un valor fijo de σ y la confianza especificada.
- Cuando σ se incrementa, el tamaño de la muestra n requerido se incrementa para una longitud deseada $2E$ fija y la confianza especificada.
- Cuando el nivel de confianza se incrementa, el tamaño de la muestra n requerido se incrementa para la longitud deseada $2E$ y la desviación estándar σ fijas.

Intervalos de confianza de una cola

También es posible obtener intervalos de confianza de una cola para μ haciendo $l = -\infty$ o bien $u = \infty$, y sustituyendo $z_{\alpha/2}$ con z_α.

El **intervalo de confianza superior** del $100(1-\alpha)$ por ciento para μ es

$$\mu \le u = \bar{x} + z_\alpha \sigma/\sqrt{n} \tag{8-37}$$

El **intervalo de confianza inferior** del $100(1-\alpha)$ por ciento para μ es

$$\bar{x} - z_\alpha \sigma/\sqrt{n} = l \le \mu \tag{8-38}$$

8-2.7 Método general para deducir un intervalo de confianza

Es sencillo dar un método general para encontrar un intervalo de confianza para un parámetro desconocido θ. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de n observaciones. Suponga que puede encontrarse un estadístico $g(X_1, X_2, \ldots, X_n; \theta)$ con las siguientes propiedades:

1. $g(X_1, X_2, \ldots, X_n; \theta)$ depende tanto de la muestra como de θ, y

2. la distribución de probabilidad de $g(X_1, X_2, ..., X_n; \theta)$ no depende de θ ni de algún otro parámetro desconocido.

En el caso considerado en esta sección, el parámetro $\theta = \mu$. La variable aleatoria $g(X_1, X_2, ..., X_n; \mu) = (X - \mu)/(\sigma/\sqrt{n})$ satisface las dos condiciones anteriores; depende de la muestra y de μ, y tiene una distribución normal estándar, ya que σ es conocida. Ahora deben encontrarse las constantes C_L y C_U tales que

$$P[C_L \leq g(X_1, X_2, \ldots, X_n; \theta) \leq C_U] = 1 - \alpha$$

Debido a la propiedad 2, C_L y C_U no dependen de θ. En el ejemplo considerado, $C_L = -z_{\alpha/z}$ y $C_U = z_{\alpha/z}$. Por último, es necesario manipular las desigualdades en las expresiones de probabilidad de modo que

$$P[L(X_1, X_2, \ldots, X_n) \leq \theta \leq U(X_1, X_2, \ldots, X_n)] = 1 - \alpha$$

Se obtiene así $L(X_1, X_2, ..., X_n)$ y $U(X_1, X_2, ..., X_n)$ como los límites de confianza inferior y superior que definen el intervalo de confianza del $100(1 - \alpha)$ por ciento para θ. En el ejemplo considerado, se encuentra $L(X_1, X_2, \ldots, X_n) = \overline{X} - z_{\alpha/2}\sigma/\sqrt{n}$ y $U(X_1, X_2, \ldots, X_n) = \overline{X} + z_{\alpha/2}\sigma/\sqrt{n}$.

8-2.8 Intervalos de confianza *bootstrap* (cargador inicial)

En la sección 7-3.4 se explicó la forma en que podía usarse una técnica llamada ***bootstrap*** (cargador inicial) para estimar el error estándar $\sigma_{\hat{\theta}}$, donde $\hat{\theta}$ es una estimación del parámetro θ. El *bootstrap* también puede usarse para encontrar un intervalo de confianza para el parámetro θ. Para ilustrar esto, considérese el caso en que θ es la media μ de una distribución normal con σ conocida. Entonces, el estimador de θ es $\overline{X}$. Obsérvese asimismo que $z_{\alpha/2}\sigma/\sqrt{n}$ es el percentil $100(1 - \alpha/2)$ de la distribución de $\overline{X} - \mu$, y que $z_{-\alpha/2}\sigma/\sqrt{n}$ es el percentil $100(\alpha/2)$ de esta distribución. Por lo tanto, el enunciado de probabilidad asociado con el intervalo de confianza del $100(1 - \alpha)$ por ciento puede escribirse como

$$P(\text{percentil } 100(\alpha/2) \leq \overline{X} - \mu \leq \text{percentil } 100(1 - \alpha/2)) = 1 - \alpha$$

o bien

$$P(\overline{X} - \text{percentil } 100(1 - \alpha/2) \leq \mu \leq \overline{X} - \text{percentil } 100(\alpha/2)) = 1 - \alpha$$

Esta última expresión de probabilidad implica que los límites de confianza inferior y superior del $100(1 - \alpha)$ por ciento para μ son

$$L = \overline{X} - \text{percentil } 100(1 - \alpha/2) \text{ de } \overline{X} - \mu = \overline{X} - z_{\alpha/2}\sigma/\sqrt{n}$$
$$U = \overline{X} - \text{percentil } 100(\sigma/2) \text{ de } \overline{X} - \mu = \overline{X} + z_{\alpha/2}\sigma/\sqrt{n}$$

Este resultado puede generalizarse para un parámetro θ arbitrario. Los límites de confianza del $100(1-\alpha)$ por ciento para θ son

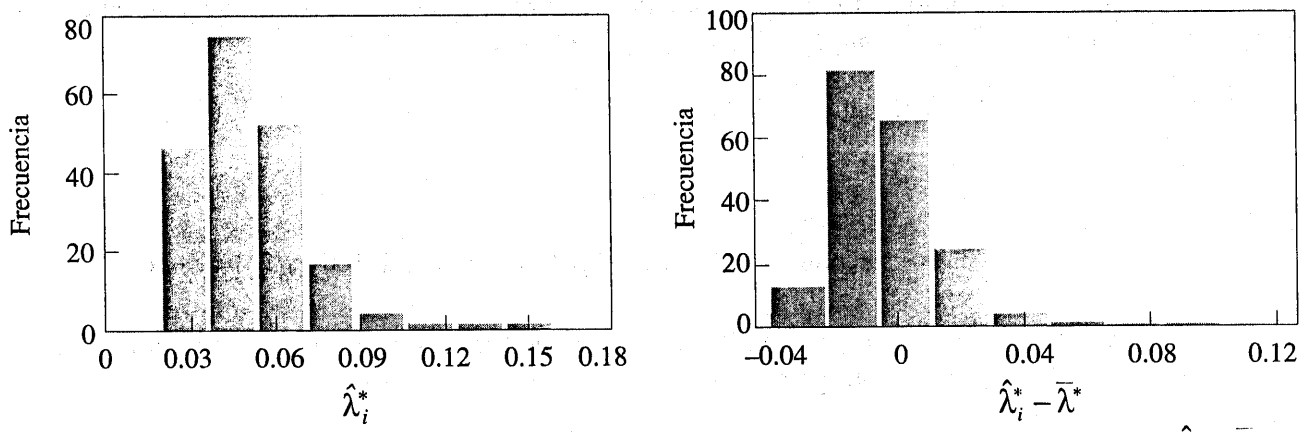
$$L = \hat{\theta} - \text{percentil } 100(1-\alpha/2) \text{ de } \hat{\theta} - \theta$$
$$U = \hat{\theta} - \text{percentil } 100(\alpha/2) \text{ de } \hat{\theta} - \theta$$

Desafortunadamente, los percentiles de $\hat{\theta} - \theta$ pueden no ser fáciles de encontrar, como en el caso de la media de la distribución normal. Sin embargo, podrían estimarse a partir de **muestras *bootstrap*.** Suponga que se encuentran B muestras *bootstrap* y que se calculan $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$ y $\bar{\theta}^*$, después se calculan $\hat{\theta}_1^* - \bar{\theta}^*, \hat{\theta}_2^* - \bar{\theta}^*, \ldots, \hat{\theta}_B^* - \bar{\theta}^*$. Los percentiles que se requieren pueden obtenerse directamente de las diferencias. Por ejemplo, si $B = 200$ y se quiere un intervalo de confianza de 95% para θ, entonces la quinta diferencia $\hat{\theta}_i^* - \bar{\theta}^*$ más pequeña y la quinta más grande son las estimaciones de los percentiles necesarios.

Se ilustrará este procedimiento utilizando la situación que se describió primero en el ejemplo 7-3, en la que interviene el parámetro λ de una distribución exponencial. Recuérdese que en ese ejemplo se probó una muestra aleatoria de $n = 8$ módulos electrónicos usados en el controlador del motor de un automóvil hasta encontrar una falla, y la estimación de λ obtenida fue = 0.0462, donde $\hat{\lambda} = 1/\bar{X}$ es un estimador de máxima verosimilitud. Se usaron 200 muestras *bootstrap* para obtener una estimación del error estándar de .

La figura 8-10*a* es un histograma de las 200 estimaciones *bootstrap* $\hat{\lambda}_1^*$, $i = 1, 2, \ldots, 200$. Obsérvese que el histograma no es simétrico, sino que está sesgado a la derecha, lo cual indica que la distribución de muestreo de $\hat{\lambda}$ también tiene la misma forma. El promedio muestral $\bar{\lambda}^* = 0.5013$ de estas estimaciones *bootstrap* se restó de cada $\hat{\lambda}_i^*$. El histograma de las diferencias $\hat{\lambda}_i^* - \bar{\lambda}^*$, $i = 1, 2, \ldots, 200$, se muestra en la figura 8-10*b*. Suponga que quiere encontrarse un intervalo de confianza de 90% para λ. El percentil 5 de las muestras *bootstrap* $\hat{\lambda}_i^* - \bar{\lambda}^*$ es –0.0228 y el percentil 95 es 0.03135. Por lo tanto, los límites de confianza bootstrap de 90% inferior y superior son

$$L = \hat{\lambda} - \text{percentil 95 de } \hat{\lambda}_i^* - \bar{\lambda}^* = 0.0462 - 0.03135 = 0.0149$$
$$U = \hat{\lambda} - \text{percentil 5 de } \hat{\lambda}_i^* - \bar{\lambda}^* = 0.0462 - (-0.0228) = 0.0690$$



a) Histograma de las estimaciones *bootstrap*

b) Histograma de las diferencias $\hat{\lambda}_i^* - \bar{\lambda}^*$

Figura 8-10 Histogramas de las estimaciones *bootstrap* de λ y las diferencias $\hat{\lambda}_i^* - \hat{\lambda}^*$ usadas para encontrar el intervalo de confianza *bootstrap*.

Por lo tanto, el intervalo de confianza *bootstrap* de 90% para λ es $0.0149 \leq \lambda \leq 0.0690$. Pero ocurre que en una distribución exponencial hay un intervalo de confianza exacto para el parámetro λ. Para los datos del ejemplo 7-3, el intervalo de confianza de 90% exacto[1] para λ es $0.0230 \leq \lambda \leq 0.0759$. Obsérvese que los dos intervalos de confianza son muy similares. La longitud del intervalo de confianza exacto es $0.0759 - 0.0230 = 0.0529$, en tanto que la longitud del intervalo de confianza *bootstrap* es $0.0690 - 0.0149 = 0.0541$, que es sólo un poco más grande. El método de percentiles para los intervalos de confianza *bootstrap* funciona bien cuando el estimador es insesgado y el error estándar de $\hat{\theta}$ es aproximadamente constante (como una función de θ). Una mejora, conocida como el método del *sesgo corregido y acelerado*, ajusta los percentiles en casos más generales. Podría aplicarse en este ejemplo (ya que $\hat{\lambda}$ es un estimador sesgado), pero a costa de una complejidad adicional.

EJERCICIOS DE LA SECCIÓN 8-2

8-19. Se requiere que la resistencia a la ruptura de la fibra textil usada en la fabricación de material para cortinas sea de al menos 100 psi. La experiencia pasada indica que la desviación estándar de la resistencia a la ruptura es 2 psi. Se prueba una muestra aleatoria de nueve observaciones, y se encuentra que la resistencia a la ruptura promedio es 98 psi.

a) ¿La fibra se considerará aceptable con $\alpha = 0.05$?

b) ¿Cuál es el valor P para esta prueba?

c) ¿Cuál es la probabilidad de aceptar la hipótesis nula con $\alpha = 0.05$ si la resistencia a la ruptura real de la fibra es 104 psi?

d) Encuentre un intervalo de confianza de dos colas de 95% para la verdadera media dela resistencia a la ruptura.

8-20. Se está estudiando el rendimiento de un proceso químico. Por experiencias previas con este proceso, se sabe que la desviación estándar del rendimiento es 3. En los últimos cinco días de operación se observaron los siguientes rendimientos: 91.6%, 88.75%, 90.8%, 89.95% y 91.3%. Use $\alpha = 0.05$.

a) ¿Hay evidencia de que el rendimiento no es de 90%?

b) ¿Cuál es el valor P para esta prueba?

c) ¿Qué tamaño de la muestra se necesitaría para detectar un rendimiento medio real de 85% con probabilidad 0.95?

d) ¿Cuál es la probabilidad del error tipo II si el rendimiento medio real es 92%?

e) Encuentre un intervalo de confianza de dos colas de 95% para el rendimiento medio real.

8-21. Se sabe que el diámetro de los agujeros de un arnés para cables tiene una desviación estándar de 0.01 pulgadas. Una muestra aleatoria de tamaño 10 produce un diámetro promedio de 1.5045 pulgadas. Use $\alpha = 0.01$.

a) Pruebe la hipótesis de que la verdadera media del diámetro de los agujeros es igual a 1.50 pulgadas.

b) ¿Cuál es el valor P para esta prueba?

c) ¿Qué tamaño de la muestra se necesitaría para detectar una verdadera media del diámetro de los agujeros de 1.505 pulgadas con una probabilidad de al menos 0.90?

d) ¿Cuál es el error β si la verdadera media del diámetro de los agujeros es 1.505 pulgadas?

e) Encuentre un intervalo de confianza de dos colas de 99% para la media del diámetro de los agujeros. ¿Los resultados de este cálcu-

---

[1] El intervalo de confianza es $\chi^2_{\alpha/2,2n}/2\sum x_i \leq \lambda \leq \chi^2_{1-\alpha/2,2n}/2\sum x_i$, donde $\chi^2_{\alpha/2,2n}$ y $\chi^2_{1-\alpha/2,2n}$ son los puntos porcentuales $\alpha/2$ inferior y superior de la distribución ji-cuadrada (la cual se introducirá en la sección 8-4), y las x_i son las n observaciones muestrales.

lo parecen intuitivamente correctos con base en la respuesta de los incisos a) y b) de este problema? Favor de comentar la respuesta.

8-22. Un fabricante produce anillos para pistones de motor de automóvil. Se sabe que el diámetro de los anillos tiene una distribución aproximadamente normal y que tiene una desviación estándar de $\sigma = 0.001$ mm. Una muestra aleatoria de 15 anillos tiene un diámetro medio de $\bar{x} = 74.036$ mm.

a) Pruebe la hipótesis de que el diámetro medio de los anillos para pistones es 74.035 mm. Use $\alpha = 0.01$.

b) ¿Cuál es el valor P para esta prueba?

c) Construya un intervalo de confianza de dos colas de 99% para el diámetro medio de los anillos para pistones.

d) Construya un límite de confianza inferior de 95% para el diámetro medio de los anillos para pistones.

8-23. Se sabe que la vida en horas de una bombilla de 75 watts tiene una distribución aproximadamente normal, con desviación estándar $\sigma = 5$ horas. Una muestra aleatoria de 20 bombillas tiene una vida media de $\bar{x} = 1\ 014$ horas.

a) ¿Hay evidencia que apoye la afirmación de que la vida de las bombillas excede 1 000 horas? Use $\alpha = 0.05$.

b) ¿Cuál es el valor P para la prueba en el inciso a)?

c) ¿Cuál es el error β para la prueba del inciso a) si la verdadera vida media es 1 050 horas?

d) ¿Qué tamaño de la muestra se necesitaría para asegurarse de que β no excede 0.10 si la verdadera vida media es 1 025 horas?

e) Construya un intervalo de confianza de dos colas de 95% para la vida media.

f) Construya un intervalo de confianza inferior de 95% para la vida media.

8-24. Un ingeniero civil analiza la fuerza de compresión del concreto. La fuerza de compresión tiene una distribución aproximadamente normal con una varianza de $\sigma^2 = 1\ 000\ (\text{psi})^2$. Una muestra aleatoria de 12 observaciones tiene una media de la fuerza de compresión de $\bar{x} = 3\ 250$ psi.

a) Pruebe la hipótesis de que la media de la fuerza de compresión es 3 500 psi. Use $\alpha = 0.01$.

b) ¿Cuál es el nivel de significación más bajo en el que usted estaría dispuesto a rechazar la hipótesis nula?

c) Construya un intervalo de confianza de dos colas de 95% para la media de la fuerza de compresión.

d) Construya un intervalo de confianza de dos colas de 99% para la media de la fuerza de compresión. Compare la longitud de este intervalo de confianza con la del intervalo que se encontró en el inciso c).

8-25. Suponga que en el ejercicio 8-23 quería tenerse una confianza de 95% de que el error en la estimación de la vida media fuera menor que 5 horas. ¿Qué tamaño de la muestra deberá usarse?

8-26. Suponga que en el ejercicio 8-23 se quería que la longitud total del intervalo de confianza de dos colas para la vida media fuera de 6 horas con una confianza de 95%. ¿Qué tamaño de la muestra deberá usarse?

8-27. Suponga que en el ejercicio 8-24 se desea estimar la fuerza de compresión con un error que sea menor que 15 psi con una confianza de 99%. ¿Qué tamaño de la muestra se necesita?

8-3 INFERENCIA SOBRE LA MEDIA DE UNA POBLACIÓN, VARIANZA DESCONOCIDA

Cuando se prueban hipótesis o se construyen intervalos de confianza para la media μ de una población cuando no se conoce σ^2, pueden usarse los procedimientos de prueba de la sección

8-2, siempre que el tamaño de la muestra sea grande (digamos $n \geq 30$). Estos procedimientos son aproximadamente válidos (debido al teorema del límite central) independientemente de si la población fundamental es normal o no. Sin embargo, cuando la muestra es pequeña y no se conoce σ^2, debe hacerse un supuesto acerca de la forma de la distribución fundamental a fin de obtener un procedimiento de prueba. Un supuesto razonable en muchos casos es que la distribución fundamental es normal.

La distribución normal es una buena aproximación de muchas poblaciones que se encuentran en la práctica, por lo que este supuesto llevará a procedimientos de inferencia que pueden aplicarse en múltiples situaciones. De hecho, una desviación moderada de la normalidad afectará poco la validez. Cuando el supuesto no es razonable, una alternativa es usar procedimientos no paramétricos que son válidos para cualquier distribución fundamental. Véase el capítulo 13 para una introducción de estas técnicas.

8-3.1 Pruebas de hipótesis sobre la media

Suponga que la población de interés tiene una distribución normal con media μ y varianza σ^2 desconocidas. Quiere probarse la hipótesis de que μ es igual a una constante μ_0. Obsérvese que esta situación es similar a la de la sección 8-2, excepto porque ahora no se conocen ni μ ni σ^2. Suponga que se cuenta con una muestra aleatoria de tamaño n, digamos $X_1, X_2, \ldots, X_n$, y sean $\bar{X}$ y S^2 la media y la varianza muestrales, respectivamente.

Quiere probarse la hipótesis alternativa de dos colas

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Si se conoce la varianza σ^2, el estadístico de la prueba es la ecuación 8-10:

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Cuando no se conoce σ^2, un procedimiento lógico es reemplazar σ en la ecuación 8-10 con la desviación estándar muestral S. Entonces, el estadístico de la prueba es:

$$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \tag{8-39}$$

Una pregunta lógica es, ¿qué efecto tiene sobre la distribución del estadístico T_0 reemplazar σ con S? Si n es grande, la respuesta a esta pregunta es "muy poco", y puede procederse a aplicar el procedimiento de prueba basado en la distribución normal de la sección 8-2. Sin embargo, n por lo general es pequeña en la mayoría de los problemas de ingeniería, y en este caso debe emplearse una distribución diferente.

Definición

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria para una distribución normal con media μ desconocida y varianza σ^2 desconocida. La cantidad

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

tiene una distribución t con $n - 1$ grados de libertad.

La función de densidad de probabilidad t es

$$f(x) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\,\Gamma(k/2)} \cdot \frac{1}{[(x^2/k)+1]^{(k+1)/2}} \quad -\infty < x < \infty \tag{8-40}$$

donde k es el número de grados de libertad. La media y la varianza de la distribución t son cero y $k/(k-2)$ (para $k > 2$), respectivamente.

En la figura 8-11 se presentan varias distribuciones t. La apariencia general de la distribución t es similar a la de la distribución normal estándar, por cuanto ambas distribuciones son simétricas y unimodales, y el valor máximo de las ordenadas se alcanza cuando la media $\mu = 0$. Sin embargo, la distribución t tiene colas de mayor peso que la normal; es decir, tiene más probabilidad en las colas que la distribución normal. Cuando el número de grados de libertad $k \to \infty$, la forma límite de la distribución t es la distribución normal estándar. Al visualizar la distribución t, en ocasiones es útil saber que la ordenada de la densidad en la media $\mu = 0$ es aproximadamente cuatro o cinco veces mayor que la ordenada en los percentiles 5 y 95. Por ejemplo, con 10 grados de libertad para t esta relación es 4.8, con 20 grados de libertad es 4.3, y con 30 grados de libertad, 4.1. Por comparación, este factor es 3.9 para la distribución normal.

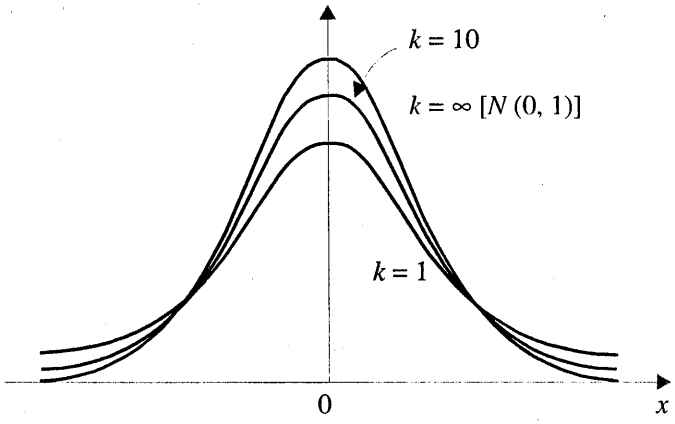


Figura 8-11 Funciones de densidad de probabilidad de varias distribuciones t.

En la tabla IV del apéndice se presentan los **puntos porcentuales** de la distribución t. Se hará que $t_{\alpha, k}$ sea el valor de la variable aleatoria T con k grados de libertad arriba del cual se encuentra un área (o probabilidad) α. Por tanto, $t_{\alpha, k}$ es un punto porcentual 100α de la cola superior de la distribución t con k grados de libertad. En la figura 8-12 se ilustra este punto porcentual. En la tabla IV del apéndice, los valores α son los encabezados de las columnas, y los grados de libertad se enlistan en la columna izquierda. Para ilustrar el uso de la tabla, obsérvese que el valor t con 10 grados de libertad que tiene un área de 0.05 a la derecha es $t_{0.05, 10} = 1.812$. Es decir,

$$P(T_{10} > t_{0.05,10}) = P(T_{10} > 1.812) = 0.05$$

Puesto que la distribución t es simétrica alrededor de cero, se tiene $t_{1-\alpha} = -t_{\alpha}$; es decir, el valor t que tiene un área de $1 - \alpha$ a la derecha (y por lo tanto un área de α a la izquierda) es igual al negativo del valor t que tiene área α en la cola derecha de la distribución. Por lo tanto, $t_{0.95, 10} = -t_{0.05, 10} = -1.812$.

Ahora puede verse con claridad que la distribución del estadístico de la prueba de la ecuación 8-39 es t con $n - 1$ grados de libertad si la hipótesis nula H_0: $\mu = \mu_0$ es verdadera. Para probar H_0: $\mu = \mu_0$, se calcula el valor del estadístico de prueba t_0 en la ecuación 8-39, y H_0 se rechaza si

$$t_0 > t_{\alpha/2, n-1} \tag{8-41a}$$

o si

$$t_0 > -t_{\alpha/2, n-1} \tag{8-41b}$$

donde $t_{\alpha/2, n-1}$ y $-t_{\alpha/2, n-1}$ son los puntos porcentuales $100\alpha/2$ superior e inferior de la distribución t con $n - 1$ grados de libertad definidos arriba.

Para la hipótesis alternativa de una cola

$$\begin{aligned} H_0&: \mu = \mu_0 \\ H_1&: \mu > \mu_0 \end{aligned} \tag{8-42}$$

se calcula el estadístico de prueba t_0, utilizando la ecuación 8-39, y H_0 se rechaza si

$$t_0 > t_{\alpha, n-1} \tag{8-43}$$

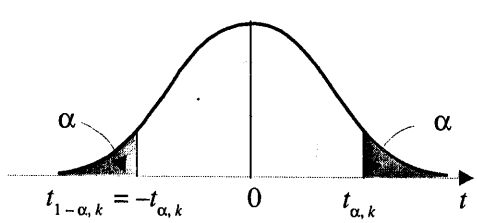


Figura 8-12 Puntos porcentuales de la distribución t.

Para la otra hipótesis alternativa de una cola

$$H_0: \mu = \mu_0$$
$$H_1: \mu < \mu_0 \tag{8-44}$$

H_0 se rechaza si

$$t_0 < -t_{\alpha, n-1} \tag{8-45}$$

EJEMPLO 8-8

Un artículo de la revista *Materials Engineering* (Vol. II, No. 4, pp. 275-281) describe los resultados de las pruebas por tracción para la adhesión en 22 muestras de la aleación U-700. Las cargas en que fallaron las muestras (en MPa) son las siguientes:

19.8	18.5	17.6	16.7	15.8
15.4	14.1	13.6	11.9	11.4
11.4	8.8	7.5	15.4	15.4
19.5	14.9	12.7	11.9	11.4
10.1	7.9			

La media muestral es $\bar{x} = 13.71$, y la desviación estándar muestral es $s = 3.55$. ¿Los datos sugieren que la carga para la falla media excede 10 MPa? Suponga que la carga para la falla tiene una distribución normal, y use $\alpha = 0.05$.

La solución usando el procedimiento de ocho pasos para la prueba de hipótesis es la siguiente:

1. El parámetro de interés es la carga para la falla media, μ.
2. $H_0: \mu = 10$
3. $H_1: \mu > 10$. Quiere rechazarse H_0 si la carga para la falla media excede 10 MPa.
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

6. H_0 se rechaza si $t_0 > t_{0.05, 21} = 1.721$.
7. Cálculos: Puesto que $\bar{x} = 13.71$, $s = 3.55$, $\mu_0 = 10$, y $n = 22$, se tiene

$$t_0 = \frac{13.71 - 10}{3.55/\sqrt{22}} = 4.90$$

8. Conclusiones: Puesto que $t_0 = 4.90 > 1.721$, se rechaza H_0 y se concluye que en el nivel de significación 0.05 la carga para la falla media excede 10 MPa.

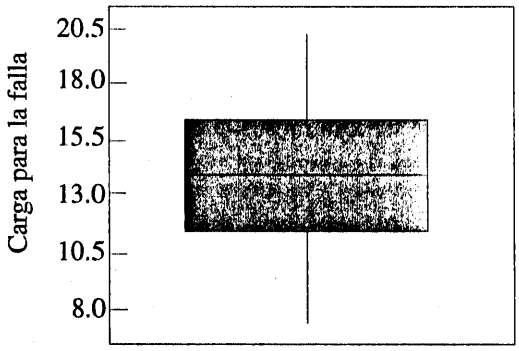


Figura 8-13 Gráfica de caja y bigote para los datos de la carga para la falla del ejemplo 8-8.

Como se señaló anteriormente, en la prueba t se supone que las observaciones son una muestra aleatoria de una población normal. Siempre es una buena idea investigar la **validez de los supuestos** cuando se aplica cualquier procedimiento estadístico. En la figura 8-13 se presenta una gráfica de caja de las 22 observaciones de la carga para la falla del ejemplo 8-8. La impresión que se obtiene al observar esta gráfica es que la muestra proviene de una población simétrica y que no hay razones inmediatas para cuestionar el supuesto de normalidad.

Otra forma excelente de verificar el supuesto de normalidad en la prueba t es examinando la gráfica de probabilidad normal de los datos muestrales. La figura 8-14 es una gráfica de probabilidad normal de los datos de la carga para la falla. Las observaciones están muy cerca de la línea recta, por lo que se concluye que el supuesto de normalidad es razonable.

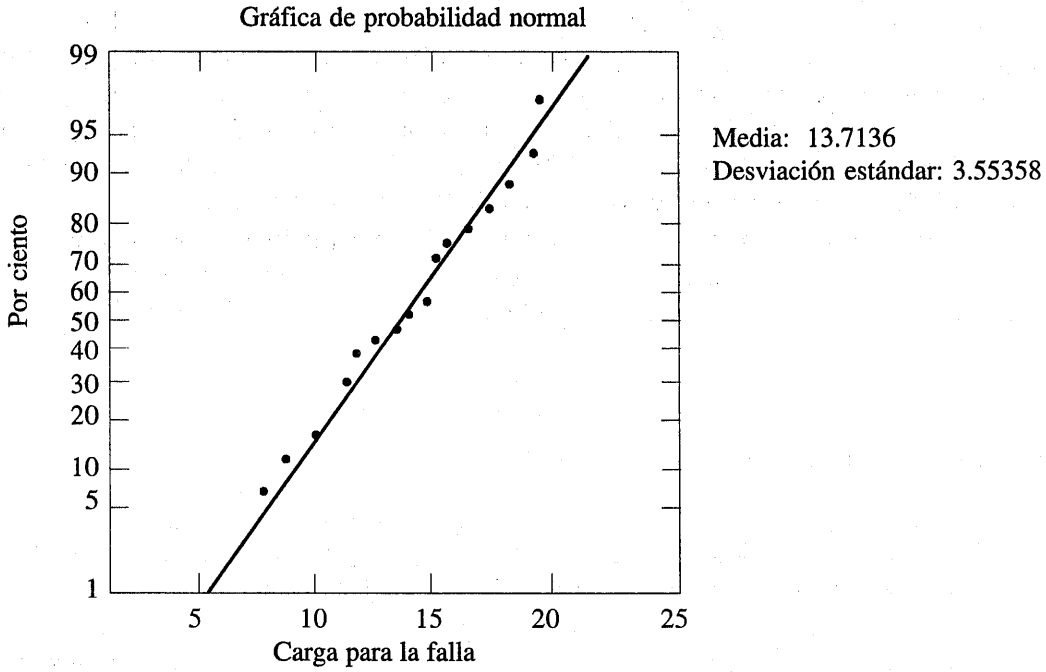


Figura 8-14 Gráfica de probabilidad normal de los datos de carga para la falla del ejemplo 8-8.

8-3.2 Valor P para una prueba t

El valor P para una prueba t es simplemente el nivel de significación más bajo en el que se rechazaría la hipótesis nula. Es decir, es el área de la cola más allá del valor del estadístico de la prueba t_0 para una prueba de una cola, o el doble de esta área para una prueba de dos colas. Puesto que la tabla t de la tabla IV del apéndice sólo contiene 10 valores críticos para cada distribución t, el cálculo directo del valor P exacto en la tabla por lo general es imposible. Sin embargo, es fácil encontrar las cotas superior e inferior del valor P utilizando esta tabla.

Para ilustrar, considérese la prueba t basada en 21 grados de libertad del ejemplo 8-8. Los valores críticos pertinentes de la tabla IV del apéndice son los siguientes:

Valor crítico:	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
Área de cola:	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005

Obsérvese que, como $t_0 = 4.90$ en el ejemplo 8-8, se tiene $P(T_{21} > 4.90) < 0.0005$, ya que $t_0 = 4.90$ es mayor que 3.819. Por lo tanto, el valor P debe ser menor que 0.0005 para esta prueba, y podría decirse que una cota superior del valor P es 0.0005.

Suponga que el valor calculado del estadístico de la prueba fue 2.75. Entonces este valor calculado está entre los dos valores críticos 2.518 (correspondiente a $\alpha = 0.01$) y 2.831 (correspondiente a 0.005). Por lo tanto, se sabe que el valor P es menor que 0.01 pero que debe ser mayor que 0.005. Se obtienen así las cotas inferior y superior del valor P como $0.005 < P < 0.01$.

El ejemplo 8-8 es una prueba de cola superior. Si la prueba es de cola inferior, simplemente se cambia el signo de t_0 y se procede como arriba para la prueba de cola superior. Por ejemplo, si $t_0 = -2.75$ para una prueba de cola inferior con 21 grados de libertad, se encontraría que las cotas inferior y superior son $0.005 < P < 0.01$, exactamente como arriba. Recuérdese que para una prueba de dos colas el nivel de significación asociado con un valor crítico particular es el doble del área de la cola correspondiente en el encabezado de la columna. Esta consideración debe tenerse presente cuando se calculan las cotas del valor P. Por ejemplo, suponga que $t_0 = 2.75$ para una hipótesis alternativa de dos colas basada en 21 grados de libertad. El valor $t_0 > 2.518$ (correspondiente a $\alpha = 0.02$) y $t_0 < 2.831$ (correspondiente a $\alpha = 0.01$), por lo que las cotas inferior y superior del valor P serían $0.01 < P < 0.02$ en este caso.

Por último, la mayoría de los programas de computadora reportan valores P junto con el valor calculado del estadístico de la prueba. Algunas computadoras de mano también tienen esta característica. En una de estas calculadoras, el valor P obtenido para el valor $t_0 = 4.90$ en el ejemplo 8-8 es 0.000038.

8-3.3 Solución de computadora

Existe una amplia variedad de paquetes de software de estadística, y la mayoría incluyen características para la prueba de hipótesis estadísticas. En la tabla 8-2 se presenta la salida del programa Minitab aplicado a los datos del ejemplo 8-8. Obsérvese que la salida incluye algunos estadísticos de resumen acerca de la muestra, así como un intervalo de confianza de 95% para la media. (El analista puede escoger el nivel de confianza.) En la sección 8-3.5 se indicará cómo se construyó este intervalo de confianza. El programa también prueba las hipótesis de interés,

Tabla 8-2 Análisis de una muestra del ejemplo 8-8

Intervalos de confianza					
Variable	*N*	Media	Desv. Est.	Media EE	95.0%CI
Carga en la falla	22	13.714	3.554	0.758	(12.138, 15.289)

Prueba *T* de la media						
Prueba de mu = 10.000 contra mu > 10.000						
Variable	*N*	Media	Desv. Est.	Media EE	*T*	*P*
Carga en la falla	22	13.714	3.554	0.758	4.90	0.0000

permitiendo al analista especificar μ_0 (en este caso 10) y la naturaleza de la hipótesis alternativa (H_1: $\mu > 10$).

La salida incluye el valor calculado de t_0 y el valor P. El valor P se redondea a la diezmilésima más próxima, lo cual da como resultado 0.0000 en este ejemplo. Obsérvese que H_0: $\mu = 10$ deberá rechazarse en favor de H_1: $\mu > 10$, que es exactamente la misma conclusión a la que se llegó en el ejemplo 8-8.

8-3.4 Elección del tamaño de la muestra

La probabilidad del error tipo II en las pruebas para la media de una distribución normal con varianza desconocida depende de la distribución del estadístico de la prueba en la ecuación 8-39 cuando la hipótesis nula H_0: $\mu = \mu_0$ es falsa. Cuando el verdadero valor de la media es $\mu = \mu_0 + \delta$, a la distribución de T_0 se le llama **distribución *t* no central** con $n - 1$ grados de libertad y parámetro de no centralidad $\delta\sqrt{n}/\sigma$. Obsérvese que si $\delta = 0$, la distribución t no central se reduce a la **distribución *t* central** usual. Por lo tanto, el error tipo II de la hipótesis alternativa de dos colas (por ejemplo) sería

$$\begin{aligned}\beta &= P\{-t_{\alpha/2,\,n-1} \le T_0 \le t_{\alpha/2,\,n-1} \mid \delta \ne 0\} \\ &= P\{-t_{\alpha/2,\,n-1} \le T_0' \le t_{\alpha/2,\,n-1}\}\end{aligned}$$

donde T_0' denota la variable aleatoria t no central. Encuentre la probabilidad β del error tipo II para la prueba t implica encontrar la probabilidad contenida entre dos puntos de la distribución t no central. Debido a que la variable aleatoria t no central tiene una función de densidad muy complicada, esta integración debe hacerse utilizando métodos numéricos.

Por fortuna, esta ingrata tarea ya se ha realizado, y los resultados se resumen en una serie de curvas OC en las gráficas VI*e*, VI*f*, VI*g* y VI*h* del apéndice que grafican β para la prueba t contra un parámetro d para diferentes tamaños de la muestra n. En las gráficas VI*e* y VI*f* se presentan curvas para hipótesis alternativas de dos colas. El factor d en la escala de las abscisas de estas gráficas se define como

$$d = \frac{|\mu - \mu_0|}{\sigma} = \frac{|\delta|}{\sigma} \tag{8-46}$$

Para la hipótesis alternativa de una cola $\mu > \mu_0$, como en la ecuación 8-42, se usan las gráficas VI*g* y VI*h* con

$$d = \frac{\mu - \mu_0}{\sigma} = \frac{\delta}{\sigma} \tag{8-47}$$

mientras que si $\mu < \mu_0$, como en la ecuación 8-44,

$$d = \frac{\mu_0 - \mu}{\sigma} = \frac{\delta}{\sigma} \tag{8-48}$$

Se observa que d depende del parámetro desconocido σ^2. Esta dificultad puede evitarse de varias maneras. En algunos casos, pueden usarse los resultados de un experimento anterior, o información previa, para hacer una estimación inicial aproximada de σ^2. Si el interés está en evaluar el desempeño de la prueba después de que se han recabado los datos, podría usarse la varianza muestral s^2 para estimar σ^2. Si no hay experiencia previa con base en la cual estimar σ^2, entonces se define la diferencia en d media que quiere detectarse en relación con σ. Por ejemplo, si se quiere detectar una diferencia pequeña en la media, podría usarse un valor de $d = |\delta|/\sigma \leq 1$ (por ejemplo), mientras que si el interés estuviera en detectar tan sólo las diferencias moderadamente grandes en la media, podría seleccionarse $d = |\delta|/\sigma = 2$ (por ejemplo). Es decir, es el valor del cociente $|\delta|/\sigma$ lo que es importante para determinar el tamaño de la muestra, y si es posible especificar el tamaño relativo de la diferencia en las medias que quiere detectarse, entonces por lo general puede seleccionarse un valor apropiado de d.

EJEMPLO 8-9

Considérese el problema de la prueba por tracción para la adhesión del ejemplo 8-8. Si la carga para la falla media difiere de 10 MPa a lo más en 1 MPa, ¿el tamaño de la muestra $n = 22$ es adecuado para asegurar que H_0: $\mu = 10$ se rechazará con una probabilidad de al menos 0.8?

Para resolver este problema se usará la desviación estándar muestral $s = 3.55$ para estimar σ. Entonces $d = |\delta|/\sigma = 1.0/3.55 = 0.28$. Al consultar las curvas de operación característica de la gráfica VI*g* del apéndice (para $\alpha = 0.05$) con $d = 1/3.55 = 0.28$ y $n = 22$, se encuentra que $\beta = 0.68$, aproximadamente. Por tanto, la probabilidad de rechazar H_0: $\mu = 10$ si la verdadera media excede este valor en 1.0 MPa es de aproximadamente $1 - \beta = 1 - 0.68 = 0.32$, y se concluye que el tamaño de la muestra $n = 22$ no es adecuado para proporcionar la sensibilidad deseada. Para encontrar el tamaño de la muestra que se requiere para dar el grado deseado de sensibilidad, se consultan las curvas de operación característica de la gráfica VI*c* con $d = 0.28$ y $\beta = 0.2$, y se lee el tamaño de la muestra correspondiente, que es $n = 75$.

8-3.5 Intervalo de confianza para la media

Es sencillo encontrar un intervalo de confianza del $100(1 - \alpha)$ por ciento para la media de una distribución normal con varianza desconocida procediendo como se hizo en la sección 8-2.6. En general, la distribución $T = (\bar{X} - \mu)/(S/\sqrt{n})$ es t con $n - 1$ grados de libertad. Al hacer $t_{\alpha/2,\, n-1}$

el punto porcentual $100\alpha/2$ superior de la distribución t con $n - 1$ grados de libertad, puede escribirse

$$P\left(-t_{\alpha/2,n-1} \le T \le t_{\alpha/2,n-1}\right) = 1-\alpha$$

o bien

$$P\left(-t_{\alpha/2,\,n-1} \le \frac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{\alpha/2,\,n-1}\right) = 1-\alpha$$

Al reordenar esta última ecuación, se obtiene

$$P\left(\bar{X} - t_{\alpha/2,\,n-1}S/\sqrt{n} \le \mu \le \bar{X} + t_{\alpha/2,\,n-1}S/\sqrt{n}\right) = 1-\alpha \qquad (8\text{-}49)$$

Esto lleva a la siguiente definición del intervalo de confianza de dos colas del $100(1 - \alpha)$ por ciento para μ.

Definición: Intervalo de confianza para la media de una distribución normal, varianza desconocida

Si $\bar{x}$ y s son la media y la desviación estándar de una muestra aleatoria de una distribución normal con varianza σ^2 desconocida, entonces un intervalo de confianza del $100(1 - \alpha)$ por ciento para μ está dado por

$$\bar{x} = t_{\alpha/2,\,n-1}s/\sqrt{n} \le \mu \le \bar{x} + t_{\alpha/2,\,n-1}s/\sqrt{n} \qquad (8\text{-}50)$$

donde $t_{\alpha/2,\,n-1}$ es el punto porcentual $100\alpha/2$ superior de la distribución t con $n - 1$ grados de libertad.

Los intervalos de confianza de una cola para la media de una distribución normal también son motivo de interés y son fáciles de encontrar. Sólo es necesario usar el límite de confianza inferior o superior apropiado de (8-50) y reemplazar $t_{\alpha/2,\,n-1}$ con $t_{\alpha,\,n-1}$.

EJEMPLO 8-10

Considere nuevamente el problema de la prueba por tracción para la adherencia del ejemplo 8-8. Se sabe que $n = 22$, $\bar{x} = 13.71$ y $s = 3.55$. Se encontrará un intervalo de confianza de 95% para μ. Por la ecuación 8-50 se encuentra ($t_{\alpha/2,\,n-1} = t_{0.025,\,21} = 2.080$):

$$\bar{x} - t_{\alpha/2,\,n-1}s/\sqrt{n} \le \mu \le \bar{x} + t_{\alpha/2,\,n-1}s/\sqrt{n}$$

$$13.71 - 2.080(3.55)/\sqrt{22} \le \mu \le 13.71 + 2.080(3.55)/\sqrt{22}$$

$$13.71 - 1.57 \le \mu \le 13.71 + 1.57$$

$$12.14 \le \mu \le 15.28$$

En el ejemplo 8-8 se probó una hipótesis alternativa de una cola para μ. Algunos ingenieros podrían tener interés en un intervalo de confianza de una cola. El intervalo de confianza inferior de 95% para la carga de la falla media se encuentra usando el límite de confianza inferior de (8-50) con $t_{\alpha, n-1}$ en lugar de $t_{\alpha/2, n-1}$. Se obtiene así:

$$\bar{X} - t_{0.05, n-1} S/\sqrt{n} \leq \mu$$
$$13.71 - 1.721(3.25)/\sqrt{22} \leq \mu$$
$$12.41 \leq \mu$$

Por tanto, puede afirmarse con 95% de confianza que la carga de falla media excede 12.41 MPa.

EJERCICIOS DE LA SECCIÓN 8-3

8-28. El ingeniero de desarrollo de un fabricante de llantas está investigando la vida de las llantas para un nuevo compuesto de hule. Ha hecho 16 llantas y las ha probado hasta el fin de su vida útil en una prueba de carretera. La media y la desviación estándar muestrales son 60 139.7 y 3 645.94 km.

a) Al ingeniero le gustaría demostrar que la vida media de esta nueva llanta excede 60 000 km. Formule y pruebe las hipótesis apropiadas, y establezca conclusiones usando $\alpha = 0.05$.

b) Suponga que si la vida media es hasta de 61 000 km, al ingeniero le gustaría detectar esta diferencia con una probabilidad de al menos 0.90. ¿Fue adecuado el tamaño de la muestra $n = 16$ usado en el inciso a)? Use la desviación estándar muestral s como estimación de σ para llegar a una decisión.

c) Encuentre un intervalo de confianza de 95% para la vida media de las llantas.

8-29. Se efectuó una prueba de impacto Izod en 20 muestras de tubería PVC. La norma ASTM para este material requiere que la resistencia al impacto Izod debe ser mayor que 1.0 libras-pie/pulgada. El promedio y la desviación estándar muestrales obtenidos fueron $\bar{x} = 1.25$ y $s = 0.25$, respectivamente. Pruebe H_0: $\mu = 1.0$ contra H_1: $\mu > 1.0$ usando $\alpha = 0.01$, y establezca conclusiones. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

8-30. La brillantez de un cinescopio de televisión puede evaluarse midiendo la cantidad de corriente que se necesita para alcanzar un nivel de brillantez particular. Un ingeniero ha diseñado un cinescopio que en su opinión necesitará 300 microamperes de corriente para producir el nivel de brillantez deseado. Una muestra de 10 cinescopios da como resultado $\bar{x} = 317.2$ y $s = 15.7$. Formule y pruebe una hipótesis apropiada usando $\alpha = 0.05$. Encuentre el valor P para esta prueba. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

8-31. Se analizó una marca particular de margarina dietética para determinar el nivel de ácido graso polinsaturado (en porcentaje). Una muestra de seis paquetes dio como resultado los siguientes datos: 16.8, 17.2, 17.4, 16.9, 16.5, 17.1.

a) Pruebe la hipótesis H_0: $\mu = 17.0$ contra H_1: $\mu \neq 17.0$ usando $\alpha = 0.01$. ¿A qué conclusiones se llega? Use una gráfica de probabilidad normal para probar el supuesto de normalidad.

b) Encuentre el valor P para la prueba del inciso a).

c) Suponga que si la media del contenido de ácido graso polinsaturado es en realidad $\mu = 17.5$, es importante detectar este he-

cho con una probabilidad de al menos 0.90. ¿El tamaño de la muestra $n = 6$ es adecuado? Use la desviación estándar muestral para estimar la desviación estándar poblacional σ. Use $\alpha = 0.01$.

d) Encuentre un intervalo de confianza de 99% para la media μ. Ofrezca una interpretación práctica de este intervalo.

8-32. Un ingeniero civil está probando la resistencia de compresión del concreto. Prueba 12 muestras y obtiene los siguientes datos:

2 216	2 237	2 249	2 204
2 225	2 301	2 281	2 263
2 318	2 255	2 275	2 295

a) ¿Hay alguna razón para pensar que la resistencia de compresión no tiene una distribución normal? ¿Esta serie de datos apoya su punto de vista? Incluya una gráfica en la respuesta.

b) Pruebe la hipótesis H_0: $\mu = 2\,250$ psi contra H_1: $\mu \neq 2\,250$ psi usando $\alpha = 0.05$. Saque conclusiones con base en el resultado de esta prueba.

c) Construya un intervalo de confianza de dos colas de 95% para la resistencia media.

d) Construya un intervalo de confianza inferior de 95% para la resistencia media.

8-33. Una máquina produce varillas metálicas usadas en el sistema de suspensión de un automóvil. Se selecciona una muestra aleatoria de 15 varillas, y se mide el diámetro. Los datos resultantes se muestran a continuación:

8.24 mm	8.23 mm	8.20 mm
8.21	8.20	8.28
8.23	8.26	8.24
8.25	8.19	8.25
8.26	8.23	8.24

a) Verifique el supuesto de normalidad del diámetro de las varillas.

b) ¿Hay evidencia sólida que indique que el diámetro medio de las varillas excede 8.20 mm usando $\alpha = 0.05$?

c) Encuentre el valor P para la prueba estadística realizada en el inciso b).

d) Encuentre un intervalo de confianza de dos colas de 95% para el diámetro medio de las varillas.

8-34. Un ingeniero de control de calidad midió el espesor de las paredes de 25 botellas de vidrio de 2 litros. La media muestral fue $\bar{x} = 4.05$ mm, y la desviación estándar muestral fue $s = 0.08$ mm.

a) Suponga que es importante demostrar que el espesor de las paredes de las botellas excede 4.0 mm. Formule y pruebe una hipótesis apropiada usando estos datos. Saque conclusiones con $\alpha = 0.05$. Calcule el valor P para esta prueba.

b) Encuentre un intervalo de confianza de 95% para la media del espesor de las paredes de las botellas. Interprete el intervalo obtenido.

8-35. Un artículo de *Nuclear Engineering International* describe varias características de las varillas de combustible usadas en el reactor de una compañía de energía eléctrica de Noruega. Las mediciones reportadas del enriquecimiento porcentual de 12 varillas fueron las siguientes:

2.94	2.75	2.75	2.81
2.90	2.90	2.82	2.95
3.00	2.95	3.00	3.05

a) Use una gráfica de probabilidad normal para verificar el supuesto de normalidad.

b) Pruebe la hipótesis H_0: $\mu = 2.95$ contra H_1: $\mu \neq 2.95$ usando $\alpha = 0.05$, y saque las conclusiones adecuadas. Calcule el valor P para esta prueba.

c) Encuentre un intervalo de confianza de dos colas de 99% para la media del enriquecimiento porcentual. ¿Se siente usted cómodo con la afirmación de que la media del enriquecimiento porcentual es 2.95%? ¿Por qué?

8-36. Una máquina de bebidas postmezcladas se ajusta para liberar cierta cantidad de jarabe en una cámara donde se mezcla con agua carbonatada. Se encontró que una muestra aleatoria de 25 bebidas tuvo un contenido

medio de jarabe de $\bar{x} = 1.10$ onzas líquidas y una desviación estándar de $s = 0.015$ onzas líquidas.

a) ¿Los datos presentados en este ejercicio apoyan la afirmación de que la cantidad media de jarabe servida no es 1.0 onzas líquidas? Pruebe esta afirmación usando $\alpha = 0.05$.

b) ¿Los datos presentados apoyan la afirmación de que la cantidad media de jarabe servida excede 1.0 onzas líquidas? Pruebe esta afirmación usando $\alpha = 0.05$.

c) Considere la prueba de la hipótesis del inciso a). Si la cantidad media de jarabe servida difiere de $\mu = 0.10$ hasta por 0.05, es importante detectar este hecho con una probabilidad más alta (digamos al menos 0.90). Usando s como estimación de σ, ¿qué puede decirse acerca de la adecuación del tamaño de la muestra $n = 25$ usado por los experimentadores?

d) Encuentre un intervalo de confianza de 95% para la cantidad media de jarabe servida.

8-37. Un artículo en *Journal of Composite Materials* (Vol. 23, p. 1200) describe el efecto de delaminación sobre la frecuencia propia de vigas hechas con laminados mixtos. Cinco de estas vigas delaminadas se sometieron a cargas, y las frecuencias resultantes (en Hz) fueron las siguientes:

230.66, 233.05, 232.58, 229.48, 232.58

Encuentre un intervalo de confianza de dos colas de 90% para la frecuencia propia media. ¿Los resultados de los cálculos apoyan la afirmación de que la frecuencia propia media es 235 Hz? Favor de comentar los resultados.

8-4 INFERENCIA SOBRE LA VARIANZA DE UNA POBLACIÓN NORMAL

Con frecuencia se necesitan pruebas de hipótesis e intervalos de confianza para la varianza o la desviación estándar de la población. Si se tiene una muestra aleatoria $X_1, X_2, ..., X_n$, la varianza muestral S^2 es una estimación puntual insesgada de σ^2. Cuando la población se modela con una distribución normal, se aplican las pruebas y los intervalos descritos en esta sección.

8-4.1 Pruebas de hipótesis sobre la varianza de una población normal

Suponga que quiere probarse la hipótesis de que la varianza σ^2 de una población normal es igual a un valor específico, digamos σ_0^2. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de n observaciones de esta población. Para probar

$$H_0 : \sigma^2 = \sigma_0^2$$

$$H_1 : \sigma^2 \neq \sigma_0^2 \tag{8-51}$$

se usará el estadístico de la prueba:

$$X_0^2 = \frac{(n-1)S^2}{\sigma_0^2} \tag{8-52}$$

Para definir el procedimiento de prueba será necesario conocer la distribución del estadístico de la prueba X_0^2 en la ecuación 8-52 cuando la hipótesis nula es verdadera. Ahora bien, si se está haciendo un muestreo de una distribución normal con media μ desconocida y varianza σ^2 desconocida, entonces la cantidad

$$X^2 = \frac{(n-1)S^2}{\sigma^2}$$

tiene una distribución ji-cuadrada con $n - 1$ grados de libertad, la cual se abrevia X_{n-1}^2. La función de densidad de probabilidad de una variable aleatoria ji-cuadrada es

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} \qquad x > 0 \tag{8-53}$$

donde k es el número de grados de libertad.

La media y la varianza de la distribución χ^2 son

$$\mu = k \tag{8-54}$$

y

$$\sigma^2 = 2k \tag{8-55}$$

En la figura 8-15 se presentan varias distribuciones ji-cuadrada. Obsérvese que la variable aleatoria ji-cuadrada es no negativa y que la distribución de probabilidad está sesgada a la derecha. Sin

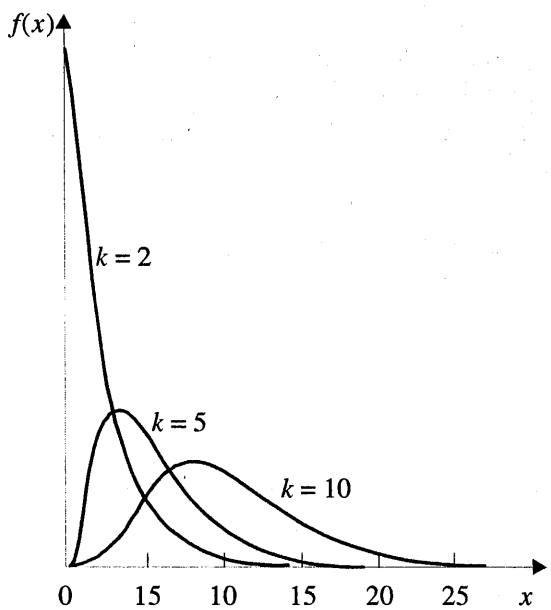


Figura 8-15 Funciones de densidad de probabilidad de varias distribuciones χ^2.

embargo, la distribución se hace más simétrica cuando k se incrementa. Cuando $k \to \infty$, la forma límite de la distribución ji-cuadrada es la distribución normal.

En la tabla III del apéndice se dan los **puntos porcentuales** de la distribución χ^2. Se define $\chi^2_{\alpha,k}$ como el punto porcentual, o valor de la variable aleatoria ji-cuadrada con k grados de libertad, tal que la probabilidad de que X^2 exceda este valor es α. Es decir,

$$P(X^2 > \chi^2_{\alpha,k}) = \int_{\chi^2_{\alpha,k}}^{\infty} f(u)du = \alpha$$

El área sombreada de la figura 8-16 indica esta probabilidad. Para ilustrar el uso de la tabla III, obsérvese que las áreas α son los encabezados de las columnas y los grados de libertad k se dan en la columna izquierda. Por lo tanto, el valor con 10 grados de libertad que tiene un área (probabilidad) de 0.05 a la derecha es $\chi^2_{0.05,10} = 18.31$. A este valor se le llama con frecuencia un punto superior de 5% de ji-cuadrada con 10 grados de libertad. Esto puede expresarse como un enunciado de probabilidad de la siguiente manera:

$$P(X^2 > \chi^2_{0.05,10}) = P(X^2 > 18.31) = 0.05$$

Es relativamente sencillo construir una prueba para la hipótesis de la ecuación 8-55. Si la hipótesis nula H_0: $\sigma^2 = \sigma_0^2$ es verdadera, entonces el estadístico de la prueba X_0^2 definido en la ecuación 8-52 sigue la distribución ji-cuadrada con $n - 1$ grados de libertad. Por lo tanto, se calcula el valor del estadístico de prueba χ_0^2, y la hipótesis H_0: $\sigma^2 = \sigma_0^2$ se rechazaría si

$$\chi_0^2 > \chi^2_{\alpha/2,n-1} \tag{8-56a}$$

o si

$$\chi_0^2 > \chi^2_{1-\alpha/2,n-1} \tag{8-56b}$$

donde $\chi^2_{\alpha/2,n-1}$ y $\chi^2_{1-\alpha/2,n-1}$ son los puntos porcentuales $100\alpha/2$ superior e inferior de la distribución ji-cuadrada con $n - 1$ grados de libertad, respectivamente.

Se usa el mismo estadístico de la prueba para las hipótesis alternativas. Para la hipótesis de una cola

$$\begin{aligned} H_0 &: \sigma^2 = \sigma_0^2 \\ H_1 &: \sigma^2 > \sigma_0^2 \end{aligned} \tag{8-57}$$

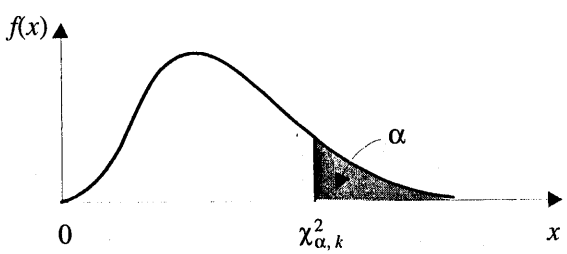


Figura 8-16 Puntos porcentuales $\chi^2_{\alpha,k}$ de la distribución χ^2.

H_0 se rechazaría si

$$\chi_0^2 > \chi_{\alpha, n-1}^2 \tag{8-58}$$

Para la otra hipótesis de una cola

$$\begin{aligned} H_0 &: \sigma^2 = \sigma_0^2 \\ H_1 &: \sigma^2 < \sigma_0^2 \end{aligned} \tag{8-59}$$

H_0 se rechazaría si

$$\chi_0^2 < \chi_{1-\alpha, n-1}^2 \tag{8-60}$$

EJEMPLO 8-11

Se usa una máquina automática para llenar botellas con un detergente líquido. Una muestra aleatoria de 20 botellas da como resultado una varianza muestral del volumen de llenado de $s^2 = 0.0153$ (onzas líquidas)^2. Si la varianza del volumen de llenado excede 0.01 (onzas líquidas)^2, una proporción inaceptable de botellas se llenarán de más o de menos. ¿Hay evidencia en los datos muestrales que sugiera que el fabricante tiene un problema con botellas llenadas de más o de menos? Use $\alpha = 0.05$, y suponga que el volumen de llenado tiene una distribución normal.

Usando el procedimiento de ocho pasos se obtiene:

1. El parámetro de interés es la varianza poblacional σ^2.
2. H_0: $\sigma^2 = 0.01$
3. H_1: $\sigma^2 > 0.01$
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

6. Se rechaza H_0 si $\chi_0^2 > \chi_{0.05,19}^2 = 30.14$.
7. Cálculos:

$$\chi_0^2 = \frac{19(0.0153)}{0.01} = 29.07$$

8. Conclusiones: Puesto que $\chi_0^2 = 29.07 < \chi_{0.05,19}^2 = 30.14$, se concluye que no hay evidencia sólida de que la varianza del volumen de llenado excede 0.01 (onzas líquidas)^2.

Usando la tabla III del apéndice, es sencillo acotar el valor P de la prueba ji-cuadrada. Por inspección de la tabla, se encuentra $\chi_{0.10,19}^2 = 27.20$ y $\chi_{0.05,19}^2 = 30.14$. Puesto que 27.20

$< 29.07 < 30.14$, se concluye que el valor P para la prueba del ejemplo 8-11 está en el intervalo $0.05 < P < 0.10$. El valor P real es $P = 0.0649$. (Este valor se obtuvo con una calculadora.)

8-4.2 Error β y elección del tamaño de la muestra

En las gáficas VI*i* a VI*n* del apéndice se presentan las curvas de operación característica para $\alpha = 0.05$ y $\alpha = 0.01$ de las pruebas ji-cuadrada de la sección 8-4.1. Para la hipótesis alternativa de dos colas de la ecuación 8-51, las gráficas VI*i* y VI*j* grafican β contra un parámetro en las abscisas

$$\lambda = \frac{\sigma}{\sigma_0} \tag{8-61}$$

para varios tamaños de la muestra n, donde σ denota el verdadero valor de la desviación estándar. Las gáficas VI*k* y VI*l* son para la hipótesis alternativa de una cola H_1: $\sigma^2 > \sigma_0^2$, en tanto que las gáficas VI*m* y VI*n* son para la otra hipótesis alternativa de una cola H_1: $\sigma^2 < \sigma_0^2$. Al utilizar estas gáficas, σ se considera como el valor de la desviación estándar que quiere detectarse.

Estas curvas pueden usarse para evaluar el error β (o potencia) asociado con una prueba particular. De manera alternativa, pueden usarse para **diseñar** una prueba; es decir, para determinar cuál es el tamaño de la muestra necesario para detectar un valor particular de σ que difiere del valor hipotético propuesto σ_0.

EJEMPLO 8-12

Considérese el problema del llenado de botellas del ejemplo 8-11. Si la varianza del proceso de llenado excede 0.01 (onzas líquidas)^2, entonces muchas botellas se llenarán de menos. Por tanto, el valor hipotético propuesto de la desviación estándar es $\sigma_0 = 0.10$. Suponga que si la verdadera desviación estándar del proceso de llenado excede este valor en 25%, habría interés en detectar este hecho con una probabilidad de al menos 0.8. ¿El tamaño de la muestra $n = 20$ es adecuado?

Para resolver este problema, se observa que se requiere que

$$\lambda = \frac{\sigma}{\sigma_0} = \frac{0.125}{0.10} = 1.25$$

Este es el parámetro de las abscisas para la carta VI*k*. A partir de esta carta, con $n = 20$ y $\lambda = 1.25$, se encuentra que $\beta \cong 0.6$. Por lo tanto, sólo hay 40% de posibilidades de que la hipótesis nula sea rechazada si la verdadera desviación estándar es en realidad hasta de $\sigma = 0.125$ onzas líquidas.

Para reducir el error β, debe usarse un tamaño de la muestra más grande. Por la curva de operación característica con $\beta = 0.20$ y $\lambda = 1.25$, se encuentra que $n = 75$, aproximadamente. Por tanto, si se quiere que el desempeño de la prueba sea como se estipuló arriba, el tamaño de la muestra debe ser de al menos 75 botellas.

8-4.3 Intervalo de confianza sobre la varianza de una población normal

En la sección anterior se señaló que si la población es normal, la distribución de muestreo de

$$X^2 = \frac{(n-1)S^2}{\sigma^2}$$

es ji-cuadrada con $n - 1$ grados de libertad. Para desarrollar el intervalo de confianza, primero se escribe

$$P(\chi^2_{1-\alpha/2,\,n-1} \le X^2 \le \chi^2_{\alpha/2,\,n-1}) = 1-\alpha$$

por lo que

$$P\left(\chi^2_{1-\alpha/2,\,n-1} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{\alpha/2,\,n-1}\right) = 1-\alpha$$

Esta última ecuación puede reordenarse como

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right) = 1-\alpha \tag{8-62}$$

Esto lleva a la siguiente definición del intervalo de confianza para σ^2.

Definición: Intervalo de confianza para la varianza de una distribución normal

Si s^2 es la varianza muestral de una muestra aleatoria de n observaciones de una distribución normal con varianza σ^2 desconocida, entonces un intervalo de confianza del $(100 - \alpha)$ por ciento para σ^2 es

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} \tag{8-63}$$

donde $\chi^2_{\alpha/2,\,n-1}$ y $\chi^2_{1-\alpha/2,\,n-1}$ son los puntos porcentuales $100\alpha/2$ superior e inferior de la distribución ji-cuadrada con $n - 1$ grados de libertad, respectivamente.

Intervalos de confianza de una cola

Para encontrar el intervalo de confianza inferior del $100(1 - \alpha)$ por ciento para σ^2, en la ecuación 8-63 el límite de confianza superior se hace igual a ∞ y se reemplaza $\chi^2_{\alpha/2,\,n-1}$ con $\chi^2_{\alpha,\,n-1}$, obteniéndose:

El intervalo de confianza inferior del $100(1 - \alpha)$ por ciento para σ^2 es

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \tag{8-64}$$

El intervalo de confianza superior del $100(1-\alpha)$ por ciento se encuentra haciendo cero el límite de confianza inferior en la ecuación 8-63 y reemplazando $\chi^2_{1-\alpha/2,\, n-1}$ con $\chi^2_{1-\alpha,\, n-1}$, obteniéndose:

El intervalo de confianza superior del $100(1-\alpha)$ por ciento para σ^2 es

$$\sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha,\, n-1}} \qquad (8\text{-}65)$$

EJEMPLO 8-13

Considérese de nuevo la máquina de llenado del ejemplo 8-11. Se sigue suponiendo que el volumen llenado tiene una distribución aproximadamente normal. Una muestra aleatoria de 20 botellas da como resultado una varianza muestral de $s^2 = 0.0153$ (onzas líquidas)^2. Se encuentra un intervalo de confianza superior de 95% a partir de la ecuación 8-63 como sigue:

$$\sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{0.95,\, 19}}$$

o bien

$$\sigma^2 \leq \frac{(19)0.0153}{10.117} = 0.0287 \quad \text{(onzas líquidas)}^2$$

Esta última expresión puede convertirse en un intervalo de confianza para la desviación estándar σ sacando raíz cuadrada en ambos miembros, de donde se obtiene

$$\sigma \leq 0.17$$

Por lo tanto, en el nivel de confianza de 95%, los datos indican que la desviación estándar del proceso podría ser hasta de 0.17 onzas líquidas.

EJERCICIOS DE LA SECCIÓN 8-4

8-38. Un remache va a insertarse en un agujero. Si la desviación estándar del diámetro del agujero excede 0.01 mm, hay una probabilidad inaceptablemente alta de que el remache no coincida con el agujero. Se selecciona una muestra aleatoria de $n = 15$ piezas y se mide el diámetro de los agujeros. La desviación estándar muestral de las mediciones del diámetro de los agujeros es $s = 0.008$ mm.

a) ¿Hay evidencia clara que indique que la desviación estándar del diámetro de los agujeros excede 0.01 mm? Use $\alpha = 0.01$. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

b) Encuentre el valor P para esta prueba.

c) Construya un intervalo de confianza inferior de 99% para σ^2.

8-39. El contenido de azúcar del jarabe en latas de duraznos en almíbar tiene una distribución normal, y se piensa que la varianza es $\sigma^2 = 18\ (\text{mg})^2$.

a) Pruebe la hipótesis H_0: $\sigma^2 = 18$ contra H_1: $\sigma^2 \neq 18$ si una muestra aleatoria de $n = 10$ latas da como resultado una desviación estándar muestral de $s = 4.8$ mg usando $\alpha = 0.05$. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

b) ¿Cuál es el valor P para esta prueba?

c) Encuentre un intervalo de confianza de dos colas de 95% para σ.

8-40. Considere los datos de la vida de las llantas del ejercicio 8-28.

a) ¿Puede concluirse, usando $\alpha = 0.05$, que la desviación estándar de la vida de las llantas excede 200 km? Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

b) Encuentre el valor P para esta prueba.

c) Encuentre un intervalo de confianza inferior de 95% para σ^2.

8-41. Considere los datos de la prueba de impacto Izod del ejercicio 8-29.

a) Pruebe la hipótesis de que $\sigma = 0.10$ contra una hipótesis alternativa que especifique que $\sigma \neq 0.10$ usando $\alpha = 0.01$, y obtenga una conclusión. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

b) ¿Cuál es el valor P para esta prueba?

c) Encuentre un intervalo de confianza de dos colas de 99% para σ^2.

8-42. Se mide el porcentaje de titanio en una aleación usada en piezas fundidas de aeronaves en 51 piezas seleccionadas al azar. La desviación estándar muestral es $s = 0.37$.

a) Pruebe la hipótesis H_0: $\sigma = 0.25$ contra H_1: $\sigma \neq 0.25$ usando $\alpha = 0.05$. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

b) Construya un intervalo de confianza de dos colas de 95% para σ.

8-43. Considere los datos del diámetro de los agujeros del ejercicio 8-38. Suponga que la verdadera desviación estándar del diámetro de los agujeros excede el valor hipotético propuesto en 50%. ¿Cuál es la probabilidad de que esta diferencia pueda detectarse en la prueba descrita en el ejercicio 8-38?

8-44. Considere los datos del contenido de azúcar del ejercicio 8-39. Suponga que la verdadera varianza es $\sigma^2 = 40$. ¿De qué tamaño necesitaría ser la muestra para detectar esta diferencia con una probabilidad de al menos 0.90?

8-5 INFERENCIA SOBRE UNA PROPORCIÓN DE UNA POBLACIÓN

Con frecuencia es necesario probar hipótesis o construir intervalos de confianza para una proporción de una población. Por ejemplo, suponga que se ha tomado una muestra aleatoria de tamaño n de una población grande (posiblemente infinita) y que X ($\leq n$) observaciones de esta muestra pertenecen a una clase de interés. Entonces $\hat{P} = X/n$ es un estimador puntual de la proporción p de la población que pertenece a esta clase. Obsérvese que n y p son los parámetros de una distribución binomial. Además, por el capítulo 3 se sabe que la distribución de muestreo de $\hat{P}$ es aproximadamente normal con media p y varianza $p(1-p)/n$, si p no está muy cerca de 0 o de 1 y si n es relativamente grande. De manera típica, para aplicar esta aproximación se necesita que np y $n(1-p)$ sean mayores o iguales que 5. En esta sección se hará uso de la aproximación normal.

8-5.1 Pruebas de hipótesis sobre una proporción binomial

En muchos problemas de ingeniería existe interés en una variable aleatoria que sigue una distribución binomial. Por ejemplo, considérese un proceso de producción en el que se fabrican artículos que se clasifican como aceptables o defectuosos. Por lo general es razonable modelar la ocurrencia de los artículos defectuosos con la distribución binomial, donde el parámetro binomial p representa la proporción de artículos defectuosos producidos. Por consiguiente, muchos problemas de toma de decisiones en la ingeniería incluyen la prueba de hipótesis acerca de p.

Se considerará la prueba de

$$\begin{aligned} H_0 &: p = p_0 \\ H_1 &: p \neq p_0 \end{aligned} \tag{8-66}$$

Se presentará una prueba aproximada basada en la aproximación normal de la distribución binomial. Como ya se señaló, este procedimiento aproximado será válido siempre que p no esté muy cerca de cero o de uno, y si el tamaño de la muestra es relativamente grande. Sea X el número de observaciones de una muestra aleatoria de tamaño n que pertenecen a la clase asociada con p. Entonces, si la hipótesis nula H_0: $p = p_0$ es verdadera, se tiene $X \sim N(np_0, np_0(1-p_0))$, aproximadamente. Para probar H_0: $p = p_0$, se calcula el **estadístico de prueba**

$$Z_0 = \frac{X - np_0}{\sqrt{np_0(1-p_0)}} \tag{8-67}$$

y H_0: $p = p_0$ se rechaza si

$$z_0 > z_{\alpha/2} \quad \text{o bien si} \quad z_0 < -z_{\alpha/2}$$

Las regiones críticas para las hipótesis alternativas de una cola se construirían de la manera usual.

EJEMPLO 8-14

Un fabricante de semiconductores produce controladores usados en dispositivos de motores de automóviles. El cliente requiere que la proporción defectuosa del proceso, en un paso crítico de la manufactura, no exceda 0.05 y que el fabricante demuestre la eficiencia del proceso en este nivel de calidad usando $\alpha = 0.05$. El fabricante de semiconductores toma una muestra aleatoria de 200 dispositivos y encuentra que cuatro de ellos están defectuosos. ¿El fabricante puede demostrarle al cliente la capacidad del proceso?

Este problema puede resolverse usando el procedimiento de prueba de hipótesis de ocho pasos como sigue:

1. El parámetro de interés es la fracción defectuosa del proceso p.
2. H_0: $p = 0.05$

3. H_1: $p < 0.05$
   Esta formulación del problema permitirá al fabricante hacer una afirmación robusta acerca de la capacidad del proceso, si se rechaza la hipótesis nula H_0: $p = 0.05$.
4. $\alpha = 0.05$
5. El estadístico de prueba es (por la ecuación 8-67)

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}}$$

   donde $x = 4$, $n = 200$, y $p_0 = 0.05$.
6. Se rechaza H_0: $p = 0.05$ si $z_0 < -z_{0.05} = -1.645$
7. Cálculos: El estadístico de prueba es

$$z_0 = \frac{4 - 200(0.05)}{\sqrt{200(0.05)(0.95)}} = -1.95$$

8. Conclusiones: Puesto que $z_0 = -1.95 < -z_{0.05} = -1.645$, se rechaza H_0 y se concluye que la fracción defectuosa del proceso p es menor que 0.05. El valor P para este valor del estadístico de la prueba z_0 es $P = 0.0256$, que es menor que $\alpha = 0.05$. Se concluye que el proceso es eficiente.

Ocasionalmente se encuentra otra forma del estadístico de prueba Z_0 de la ecuación 8-67. Nótese que si X es el número de observaciones de una muestra aleatoria de tamaño n que pertenecen a una clase de interés, entonces $\hat{P} = X/n$ es la proporción de la muestra que pertenece a esa clase. Así, al dividir tanto el numerador como el denominador de Z_0 de la ecuación 8-67 por n, se obtiene

$$Z_0 = \frac{X/n - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

o bien

$$Z_0 = \frac{\hat{P} - p_0}{\sqrt{p_0(1 - p_0)/n}} \tag{8-68}$$

Con esto, el estadístico de prueba se expresa en términos de la proporción muestral en vez del número de observaciones X de la muestra que pertenecen a la clase de interés.

8-5.2 Error tipo II y elección del tamaño de la muestra

Es posible obtener ecuaciones elementales para el error β aproximado de las pruebas de la sección 8-5.1. Suponga que p es el verdadero valor de la proporción de la población. El error β aproximado de la hipótesis alternativa de dos colas H_1: $p \neq p_0$ es

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \tag{8-69}$$

Si la hipótesis alternativa es H_1: $p < p_0$, entonces

$$\beta = 1 - \Phi\left(\frac{p_0 - p - z_{\alpha}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \tag{8-70}$$

mientras que si la hipótesis alternativa es H_1: $p > p_0$, entonces

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \tag{8-71}$$

Estas ecuaciones pueden resolverse para encontrar el tamaño de la muestra n aproximado que produce una prueba del nivel α que tiene un riesgo β especificado. Las ecuaciones para el tamaño de la muestra son

$$n = \left(\frac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_{\beta}\sqrt{p(1-p)}}{p - p_0}\right)^2 \tag{8-72}$$

para la hipótesis alternativa de dos colas y

$$n = \left(\frac{z_{\alpha}\sqrt{p_0(1-p_0)} + z_{\beta}\sqrt{p(1-p)}}{p - p_0}\right)^2 \tag{8-73}$$

para una hipótesis alternativa de una cola.

EJEMPLO 8-15

Considérese el fabricante de semiconductores del ejemplo 8-14. Suponga que la proporción defectuosa del proceso es en realidad $p = 0.03$. ¿Cuál es el error β para la prueba de la capacidad del proceso, en la que se usa $n = 200$ y $\alpha = 0.05$?

El error β puede calcularse usando la ecuación 8-70 como sigue:

$$\beta = 1 - \Phi\left(\frac{0.05 - 0.03 - (1.645)\sqrt{0.05(0.95)/200}}{\sqrt{0.03(1-0.03)/200}}\right) = 1 - \Phi(-0.44) = 0.67$$

Por tanto, la probabilidad de que el fabricante de semiconductores no pueda concluir que el proceso es eficiente es aproximadamente 0.7, si la fracción defectuosa del proceso es $p = 0.03$ (3%). Este parece ser un error β grande, pero la diferencia entre $p = 0.05$ y $p = 0.03$ es bastante pequeña, y al tamaño de la muestra $n = 200$ no es particularmente grande.

Suponga que el fabricante de semiconductores estaría dispuesto a aceptar un error β hasta de 0.10 si el verdadero valor de la fracción defectuosa del proceso fuera $p = 0.03$. Si el fabricante sigue usando $\alpha = 0.05$, ¿qué tamaño de la muestra necesitaría?

El tamaño de la muestra requerido puede calcularse con la ecuación 8-73 como sigue:

$$n = \left(\frac{1.645\sqrt{0.05(0.95)} + 1.28\sqrt{0.03(0.97)}}{0.03 - 0.05}\right)^2$$
$$\simeq 832$$

donde se ha usado $p = 0.03$ en la ecuación 8-73. Obsérvese que $n = 832$ es un tamaño de la muestra muy grande. Sin embargo, se está intentando detectar una desviación bastante pequeña del valor nulo $p_0 = 0.05$.

8-5.3 Intervalo de confianza para una proporción binomial

Puede encontrarse directamente un intervalo de confianza aproximado del $100(1 - \alpha)$ por ciento para una proporción binomial utilizando la aproximación normal. Recuérdese que la distribución de muestreo de $\hat{P}$ es aproximadamente normal con media p y varianza $p(1 - p)/n$, si p no está muy cerca de 0 o de 1 y si n es relativamente grande. Entonces, la distribución de

$$Z = \frac{X - np}{\sqrt{np(1-p)}} = \frac{\hat{P} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

es aproximadamente normal estándar.

Para construir un intervalo de confianza para p, se observa que

$$P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) \simeq 1 - \alpha$$

así que

$$P\left(-z_{\alpha/2} \le \frac{\hat{P} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \le z_{\alpha/2}\right) \simeq 1 - \alpha$$

Esta expresión puede reordenarse como

$$P\left(\hat{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le p \le \hat{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) \simeq 1-\alpha \qquad (8\text{-}74)$$

A la cantidad $\sqrt{p(1-p)/n}$ de la ecuación 8-74 se le llama el **error estándar del estimador puntual** $\hat{P}$. Desafortunadamente, los límites superior e inferior del intervalo de confianza obtenido con la ecuación 8-74 incluye el parámetro desconocido p. Sin embargo, una solución satisfactoria es reemplazar p con $\hat{P}$ en el error estándar, lo cual da como resultado

$$P\left(\hat{P} - z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \le p \le \hat{P} + z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right) \simeq 1-\alpha \qquad (8\text{-}75)$$

La ecuación 8-75 lleva al intervalo de confianza aproximado del $100(1-\alpha)$ por ciento para p.

Definición: Intervalo de confianza para una proporción

Si $\hat{p}$ es la proporción de observaciones de una muestra aleatoria de tamaño n que pertenecen a una clase de interés, entonces un intervalo de confianza del $100(1-\alpha)$ por ciento aproximado para la proporción p de la población que pertenece a esta clase es

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad (8\text{-}76)$$

donde $z_{\alpha/2}$ es el punto porcentual $\alpha/2$ superior de la distribución normal estándar.

Este procedimiento depende de la adecuación de la aproximación normal a la binomial. Siendo razonablemente conservadores, esto requiere que np y $n(1-p)$ sean mayores o iguales que 5. En situaciones donde esta aproximación no sea apropiada, particularmente en casos donde n es pequeña, deben usarse otros métodos. Podrían usarse tablas de la distribución binomial para obtener un intervalo de confianza para p. Sin embargo, preferimos el uso de métodos numéricos basados en la función de masa de probabilidad binomial que son implementados en programas de computadora.

EJEMPLO 8-16

En una muestra aleatoria de 85 rodamientos para el cigüeñal del motor de un automóvil, 10 tienen un acabado de la superficie con más asperezas de las que permiten las especificaciones. Por lo tanto, una estimación puntual de la proporción de rodamientos de la población que excede la especificación de aspereza es $\hat{p} = x/n = 10/85 = 0.12$. Con la ecuación 8-76 se calcula el intervalo de confianza de dos colas de 95%

$$\hat{p} - z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

o bien

$$0.12 - 1.96\sqrt{\frac{0.12(0.88)}{85}} \le p \le 0.12 + 1.96\sqrt{\frac{0.12(0.88)}{85}}$$

que después de simplificar queda como

$$0.05 \le p \le 0.19$$

Elección del tamaño de la muestra

Puesto que $\hat{P}$ es el estimador puntual de p, el error al estimar p con $\hat{P}$ puede definirse como $E = |p - \hat{P}|$. Obsérvese que se tiene una confianza aproximada del $100(1 - \alpha)$ por ciento de que este error sea menor que $z_{\alpha/2}\sqrt{p(1-p)/n}$. Por ejemplo, en el ejemplo 8-16 se tiene una confianza de 95% de que la proporción de la muestra $\hat{p} = 0.12$ difiere de la proporción real p en una cantidad que no excede 0.07.

En situaciones en que es posible seleccionar el tamaño de la muestra, n puede escogerse para tener una confianza del $100(1 - \alpha)$ por ciento de que el error sea menor que un valor especificado E. Si se hace $E = z_{\alpha/2}\sqrt{p(1-p)/n}$ y se resuelve la expresión para n, el tamaño de la muestra apropiado es

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) \qquad (8\text{-}77)$$

Para usar la ecuación 8-77 se necesita una estimación de p. Si se cuenta con una estimación $\hat{p}$ de una muestra anterior, ésta puede sustituir a p en la ecuación 8-77, o tal vez pueda hacerse una estimación subjetiva. Si estas alternativas no son satisfactorias, puede tomarse una muestra preliminar, calcular $\hat{p}$ y después usar la ecuación 8-77 para determinar cuántas observaciones adicionales se requieren para estimar p con la precisión deseada. Otro enfoque para elegir n utiliza el hecho de que el tamaño de la muestra de la ecuación 8-77 siempre será un máximo para $p = 0.5$ [es decir, $p(1 - p) \le 0.25$, con la igualdad para $p = 0.5$], lo cual puede usarse para encontrar una cota superior de n. En otras palabras, se tiene una confianza de al menos $100(1 - \alpha)$ por ciento de que el error al estimar p utilizando $\hat{p}$ es menor que E si el tamaño de la muestra es

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 (0.25) \qquad (8\text{-}78)$$

EJEMPLO 8-17

Considérese la situación del ejemplo 8-16. ¿Qué tamaño de la muestra se necesita si se quiere tener una confianza de 95% de que el error al usar $\hat{p}$ para estimar p es menor que 0.05? Al utilizar $\hat{p} = 0.12$ como estimación inicial de p, por la ecuación 8-77 se encuentra que el tamaño de la muestra requerido es

$$n = \left(\frac{z_{0.025}}{E}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{1.96}{0.05}\right)^2 0.12(0.88) \cong 163$$

Si se quisiera tener una confianza de *al menos* 95% de que la estimación $\hat{p}$ de la verdadera proporción p estuvo dentro de 0.05 independientemente del valor de p, entonces se usaría la ecuación 8-78 para encontrar el tamaño de la muestra

$$n = \left(\frac{z_{0.025}}{E}\right)^2 (0.25) = \left(\frac{1.96}{0.05}\right)^2 (0.25) \cong 385$$

Obsérvese que si se cuenta con información referente al valor de p, sea por una muestra anterior o por experiencias pasadas, podría usarse una muestra más pequeña conservando tanto la precisión deseada de la estimación como el nivel de confianza.

Intervalos de confianza de una cola

Es posible encontrar intervalos de confianza de una cola para aproximar p mediante una modificación simple de la ecuación 8-76.

El intervalo de confianza inferior aproximado del $100(1-\alpha)$ por ciento es

$$\hat{p} - z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \qquad (8\text{-}79)$$

El intervalo de confianza superior aproximado del $100(1-\alpha)$ por ciento es

$$p \leq \hat{p} + z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad (8\text{-}80)$$

EJERCICIOS DE LA SECCIÓN 8-5

8-45. De 1 000 casos de cáncer pulmonar seleccionados al azar, 823 resultaron en la muerte del paciente. Construya un intervalo de confianza de 95% de dos colas para el índice de decesos por cáncer pulmonar.

8-46. ¿Cuál sería el tamaño de la muestra necesario en el ejercicio 8-45 para tener una confianza de al menos 95% de que el error al estimar el índice de decesos por cáncer pulmonar sea menor que 0.03?

8-47. Una muestra aleatoria de 50 cascos de suspensión usados por motociclistas y pilotos de automóviles de carreras se sometieron a una prueba de impacto, y en 18 de estos cascos se observó algún daño.

a) Encuentre un intervalo de confianza de dos colas de 95% para la proporción real de este tipo de cascos que mostrarían daños en esta prueba.

b) Utilizando la estimación puntual de p obtenida de una muestra preliminar de 50 cascos, ¿cuántos cascos es necesario probar para tener una confianza de 95% de que el error al estimar el verdadero valor de p es menor que 0.02?

c) ¿De qué tamaño debe ser la muestra si se quiere tener una confianza de al menos 95% de que el error al estimar p es menor que 0.02, independientemente del verdadero valor de p?

8-48. El Departamento de Transporte de Arizona desea hacer una encuesta entre los residentes del estado para determinar qué proporción de la población estaría de acuerdo en aumentar el límite de velocidad en las carreteras estatales de 55 mph a 65 mph. ¿Cuántos residentes deberán encuestarse si se quiere tener una confianza de al menos 99% de que la proporción de la muestra está dentro del 0.05 de la proporción real?

8-49. Un fabricante de calculadoras electrónicas está interesado en estimar la fracción de unidades defectuosas producidas. Una muestra aleatoria de 800 calculadoras contiene 10 unidades defectuosas. Calcule un intervalo de confianza superior de 99% para la fracción defectuosa.

8-50. Va a realizarse un estudio del porcentaje de hogares que tienen al menos dos televisores. ¿De qué tamaño debe ser la muestra si se quiere tener una confianza de 99% de que el error al estimar esta cantidad es menor que 0.017?

8-51. Se estudia la fracción de circuitos integrados defectuosos producidos en un proceso de fotolitografía. Se prueba una muestra aleatoria de 300 circuitos, la cual da como resultado 13 circuitos defectuosos.

a) Use los datos para probar H_0: $p = 0.05$ contra H_1: $p \neq 0.05$. Use $\alpha = 0.05$.

b) Encuentre el valor P para esta prueba.

8-52. Considere los datos de los circuitos defectuosos del ejercicio 8-51.

a) ¿Los datos apoyan la afirmación de que la fracción de unidades defectuosas producidas es menor que 0.05 usando $\alpha = 0.05$?

b) Encuentre el valor P para esta prueba.

8-53. Un artículo de *Fortune* afirmaba que casi la mitad de todos los ingenieros continúan estudios académicos después de la licenciatura, recibiendo al final títulos de maestría o doctorado. Datos de un artículo de *Engineering Horizons* indicaban que 117 de 484 ingenieros recién graduados planeaban hacer estudios de posgrado.

a) ¿Los datos de *Engineering Horizons* son consistentes con la afirmación reportada por *Fortune*? Use $\alpha = 0.05$ y saque las conclusiones pertinentes.

b) Encuentre el valor P para esta prueba.

8-54. Un fabricante de lentes de contacto está evaluando una nueva esmeriladora mecánica. Dará el visto bueno a la máquina si el porcentaje de lentes pulidos que contienen defectos superficiales no excede 2%. Una muestra aleatoria de 250 lentes contiene seis lentes defectuosos.

a) Formule y pruebe una serie apropiada de hipótesis para determinar si puede dársele el visto bueno a la máquina. Use $\alpha = 0.05$.

b) Encuentre el valor P para la prueba del inciso a).

8-55. Un investigador afirma que al menos 10% de los cascos de futbol americano presentan defectos de fabricación que potencialmente podrían causar lesiones al jugador que los use. Una muestra de 200 cascos reveló que 16 de ellos presentaban estos defectos.

a) ¿Este descubrimiento apoya la afirmación del investigador? Use $\alpha = 0.01$.

b) Encuentre el valor P para esta prueba.

8-56. A una muestra aleatoria de 500 votantes registrados de Phoenix se les pregunta si están

en favor de usar combustibles oxigenados todo el año para reducir la contaminación ambiental. Si más de 315 votantes responden afirmativamente, se concluirá que al menos 60% de los votantes apoyan el uso de estos combustibles.

a) Encuentre la probabilidad del error tipo I si exactamente 60% de los votantes están en favor del uso de estos combustibles.

b) ¿Cuál es la probabilidad β del error tipo II si 75% de los votantes apoyan la medida?

8-57. La garantía para las baterías de teléfonos móviles se fija en 200 horas de operación, con los procedimientos de carga adecuados. Se lleva a cabo un estudio de 5 000 baterías y 15 dejan de operar antes de 200 horas. ¿Estos resultados experimentales apoyan la afirmación de que menos de 0.2% de las baterías de la compañía fallarán durante el periodo de garantía, con los procedimientos de carga adecuados? Use un procedimiento de prueba de hipótesis con $\alpha = 0.01$.

8-6 TABLA DE RESUMEN PARA PROCEDIMIENTOS DE INFERENCIA PARA UNA SOLA MUESTRA

En la tabla que aparece al principio y al final de este libro (pastas interiores), se presenta un resumen de todos los procedimientos de inferencia de una sola muestra de este capítulo. La tabla contiene el enunciado de la hipótesis nula, el estadístico de la prueba, las diferentes hipótesis alternativas y los criterios para rechazar H_0, así como las fórmulas para construir el intervalo de confianza del $100(1 - \alpha)$ por ciento.

8-7 PRUEBA DE LA BONDAD DEL AJUSTE

Los procedimientos para la prueba de hipótesis que se han revisado en las secciones anteriores están diseñados para problemas en los que la población o la distribución de probabilidad es conocida y en las hipótesis intervienen los parámetros de la distribución. Con frecuencia se encuentra otro tipo de hipótesis: no se conoce la distribución fundamental de la población, y se quiere probar la hipótesis de que una distribución particular será satisfactoria como modelo de la población. Por ejemplo, quizá quiera probarse la hipótesis de que la población es normal.

En el capítulo 5 se revisó una técnica gráfica muy útil para este problema llamada **gráfica de probabilidad** y se ilustró cómo se aplicaba en el caso de una distribución normal. En esta sección se describe un procedimiento formal para probar la bondad de ajuste basado en la distribución ji-cuadrada.

El procedimiento de prueba requiere una muestra aleatoria de tamaño n de la población cuya distribución de probabilidad es desconocida. Estas n observaciones se ordenan en un histograma de frecuencia, con k intervalos de clase. Sea O_i la frecuencia observada en el intervalo de clase i. Se calcula la frecuencia esperada a partir de la distribución de probabilidad hipotética, para el intervalo de clase i-ésimo, denotado E_i. El estadístico de prueba es

$$X_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{8-81}$$

Puede demostrarse que si la población sigue la distribución hipotética propuesta, X_0^2 tiene, aproximadamente, una distribución ji-cuadrada con $k - p - 1$ grados de libertad, donde p representa el número de parámetros de la distribución hipotética, estimados por los estadísticos muestrales. Esta aproximación mejora conforme n se incrementa. La hipótesis de que la distribución de la población es la distribución hipotética se rechazaría si el valor calculado del estadístico de prueba $\chi_0^2 > \chi_{\alpha, k-p-1}^2$.

Un punto que cabe destacar en la aplicación de este procedimiento de prueba se refiere a la magnitud de las frecuencias esperadas. Si estas frecuencias esperadas son muy pequeñas, entonces el estadístico de prueba X_0^2 no reflejará la desviación de las frecuencias observadas y las esperadas, sino únicamente la pequeña magnitud de las frecuencias esperadas. No hay consenso generalizado en cuanto al valor mínimo de las frecuencias esperadas, pero valores de 3, 4 y 5 se usan ampliamente como mínimos. Algunos autores proponen que una frecuencia esperada podría ser tan pequeña, como 1 o 2, siempre que la mayoría de ellas excedan 5. Cuando una frecuencia esperada sea muy pequeña, puede combinarse con la frecuencia esperada de un intervalo de clase adyacente. Las frecuencias observadas correspondientes también se combinarían, y k se reduciría una unidad. No es necesario que los intervalos de clase tengan la misma anchura.

Se presentan a continuación dos ejemplos del procedimiento de prueba.

EJEMPLO 8-18

Una distribución de Poisson

Se propone hipotéticamente que el número de defectos en tarjetas de circuitos impresos sigue una distribución de Poisson. Se ha colectado una muestra aleatoria de $n = 60$ tarjetas de circuitos impresos, y se observó el número de defectos. Los datos siguientes son el resultado:

Número de defectos	Frecuencia observada
0	32
1	15
2	9
3	4

La media de la distribución de Poisson supuesta en este ejemplo es desconocida y debe estimarse a partir de los datos muestrales. La estimación del número medio de defectos por tarjeta es el promedio muestral, es decir, $(32 \cdot 0 + 15 \cdot 1 + 9 \cdot 2 + 4 \cdot 3)/60 = 0.75$. En la distribución de Poisson con parámetro 0.75 puede calcularse p_i, la probabilidad teórica hipotética que se asocia con el intervalo de clase i. Puesto que cada intervalo de clase corresponde a un número particular de defectos, p_i puede encontrarse como sigue:

$$p_1 = P(X = 0) = \frac{e^{-0.75}(0.75)^0}{0!} = 0.472$$

$$p_2 = P(X = 1) = \frac{e^{-0.75}(0.75)^1}{1!} = 0.354$$

$$p_3 = P(X = 2) = \frac{e^{-0.75}(0.75)^2}{2!} = 0.133$$

$$p_4 = P(X \geq 3) = 1 - (p_1 + p_2 + p_3) = 0.041$$

Las frecuencias esperadas se calculan multiplicando el tamaño de la muestra $n = 60$ por las probabilidades p_i. Es decir, $E_i = np_i$. Las frecuencias esperadas se muestran abajo.

Número de defectos	Probabilidad	Frecuencia esperada
0	0.472	28.32
1	0.354	21.24
2	0.133	7.98
3 (o más)	0.041	2.46

Puesto que la frecuencia esperada en la última celda es menor que 3, se combinan las dos últimas celdas:

Número de defectos	Frecuencia observada	Frecuencia esperada
0	32	28.32
1	15	21.24
2 (o más)	13	10.44

El estadístico de la prueba ji-cuadrada de la ecuación 8-81 tendrá $k - p - 1 = 3 - 1 - 1 = 1$ grado de libertad, debido a que la media de la distribución de Poisson se estimó a partir de los datos.

Ahora puede aplicarse el procedimiento de prueba de hipótesis de ocho pasos, usando $\alpha = 0.05$, como sigue:

1. La variable de interés es la forma de la distribución de los defectos en las tarjetas de circuitos impresos.
2. H_0: La forma de la distribución de los defectos es de Poisson.
3. H_1: La forma de la distribución de los defectos no es de Poisson.
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(o_i - E_i)^2}{E_i}$$

6. Se rechaza H_0 si $\chi_0^2 > \chi_{0.05,1}^2 = 3.84$.

7. Cálculos:

$$\chi_0^2 = \frac{(32-28.32)^2}{28.32} + \frac{(15-21.24)^2}{21.24} + \frac{(13-10.44)^2}{10.44} = 2.94$$

8. Conclusiones: Puesto que $\chi_0^2 = 2.94 < \chi_{0.05,1}^2 = 3.84$, no puede rechazarse la hipótesis nula de que la distribución de los defectos en las tarjetas de circuitos impresos es de Poisson. El valor P para la prueba es $P = 0.0864$. (Este valor se calculó usando una calculadora HP-48.)

EJEMPLO 8-19

Una distribución continua

Un ingeniero está probando una fuente de poder usada en una computadora notebook. Utilizando $\alpha = 0.05$, él quiere determinar si una distribución normal describe adecuadamente el voltaje de salida. De una muestra aleatoria de $n = 100$ unidades obtiene las estimaciones muestrales de la media y la desviación estándar $x = 5.04$ V y $s = 0.08$ V.

Una práctica común cuando se construyen los intervalos de clase para la distribución de frecuencia usada en la prueba ji-cuadrada de la bondad del ajuste es elegir los límites de clase de las celdas de tal modo que las frecuencias esperadas $E_i = np_i$ sean iguales para todas las celdas o intervalos de clase. Para usar este método, los límites de clase $a_0, a_1, ..., a_k$ de los k intervalos de clase se elegirían de tal modo que todas las probabilidades

$$p_i = P(a_{i-1} \le X \le a_i) = \int_{a_{i-1}}^{a_i} f(x)dx$$

sean iguales. Suponga que se decide usar $k = 8$ intervalos de clase. Para la distribución normal estándar, los intervalos que dividen la escala en ocho segmentos igualmente factibles son [0, 0.32), [0.32, 0.675), [0.675, 1.15), [1.15, ∞) y los cuatro intervalos "reflejados" al otro lado de cero. Para cada intervalo $p_i = 1/8 = 0.125$, por lo que las frecuencias esperadas de las celdas son $E_i = np_i = 100(0.125) = 12.5$. La tabla completa de las frecuencias observadas y las esperadas se presenta abajo.

Intervalo de clase	Frecuencia observada o_i	Frecuencia esperada E_i
$x < 4.948$	12	12.5
$4.948 \le x < 4.986$	14	12.5
$4.986 \le x < 5.014$	12	12.5
$5.014 \le x < 5.040$	13	12.5
$5.040 \le x < 5.066$	12	12.5
$5.066 \le x < 5.094$	11	12.5
$5.094 \le x < 5.132$	12	12.5
$5.132 \le x$	14	12.5
Totales	100	100

La cota del primer intervalo de clase es $x - 1.15s = 4.948$. Para el segundo intervalo de clase es $[\bar{x} - 1.15s,\ \bar{x} - 0.675s)$, y así sucesivamente. Puede aplicarse el procedimiento de prueba de hipótesis de ocho pasos en este problema.

1. La variable de interés es la forma de la distribución del voltaje de la fuente de poder.
2. H_0: La forma de la distribución es normal.
3. H_1: La forma de la distribución no es normal.
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(o_i - E_i)^2}{E_i}$$

6. Puesto que se estimaron dos parámetros de la distribución normal, el estadístico ji-cuadrada anterior tiene $k - p - 1 = 8 - 2 - 1 = 5$ grados de libertad. Por lo tanto, se rechazará H_0 si $\chi_0^2 > \chi_{0.05,5}^2 = 11.07$.
7. Cálculos:

$$\begin{aligned}\chi_0^2 &= \sum_{i=1}^{8} \frac{(o_i - E_i)^2}{E_i} \\ &= \frac{(12-12.5)^2}{12.5} + \frac{(14-12.5)^2}{12.5} + \cdots + \frac{(14-12.5)^2}{12.5} \\ &= 0.64\end{aligned}$$

8. Conclusiones: Puesto que $\chi_0^2 = 0.64 < \chi_{0.05,5}^2 = 11.07$, no puede rechazarse H_0 y no hay evidencia robusta que indique que el voltaje de salida no tenga una distribución normal. El valor P del estadístico ji-cuadrada $\chi_0^2 = 0.64$ es $P = 0.9861$.

EJERCICIOS DE LA SECCIÓN 8-7

8-58. Considere la siguiente tabla de frecuencias de las observaciones de la variable aleatoria X:

Valores	0	1	2	3	4
Frecuencia observada	24	30	31	11	4

a) Con base en estas 100 observaciones, ¿una distribución de Poisson con una media de 1.2 es un modelo apropiado? Lleve a cabo un procedimiento de bondad del ajuste con $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

8-59. Sea que X denote el número de imperfecciones observadas en un rollo grande de acero galvanizado. Se inspeccionaron 75 rollos y se observaron los siguientes datos para los valores de X:

Valores	1	2	3	4	5	6	7	8
Frecuencia observada	1	11	8	13	11	12	10	9

a) ¿El supuesto de un distribución de Poisson parece apropiado como modelo de probabilidad para estos datos? Use $\alpha = 0.01$.

b) Calcule el valor P para esta prueba.

8-60. El número de llamadas telefónicas que entran en un conmutador desde el mediodía hasta la 1 pm en los días hábiles del lunes al

viernes se monitorea durante 6 semanas, es decir, 30 días. Sea que X se defina como el número de llamadas durante ese periodo de una hora. Se registró la frecuencia relativa de las llamadas como:

Valores	5	6	8	9	10
Frecuencia relativa	0.067	0.067	0.100	0.133	0.200
Valores	11	12	13	14	15
Frecuencia relativa	0.133	0.133	0.067	0.033	0.067

a) ¿El supuesto de una distribución de Poisson parece apropiado como modelo de probabilidad para estos datos? Use $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

8-61. Considere la siguiente tabla de frecuencias de las observaciones de la variable aleatoria X:

Valores	0	1	2	3	4
Frecuencia	4	21	10	13	2

a) Con base en estas 50 observaciones, ¿la distribución binomial con $n = 6$ y $p = 0.25$ es un modelo apropiado? Realice un procedimiento de bondad del ajuste con $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

8-62. Defina X como el número de botellas llenadas de menos en una operación de llenado de una caja de 24 botellas. Se inspeccionan 60 cajas y se registran las siguientes observaciones de X:

Valores	0	1	2	3
Frecuencia	39	23	12	1

a) Con base en estas 75 observaciones, ¿la distribución binomial es un modelo apropiado? Realice un procedimiento de bondad del ajuste con $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

8-63. El número de automóviles que circulan hacia el este por la intersección de la avenida Mill y la avenida Universidad ha sido tabulado por un grupo de estudiantes de ingeniería civil. Obtuvieron los datos siguientes:

Vehículos por minuto	Frecuencia observada	Vehículos por minuto	Frecuencia observada
40	14	53	102
41	24	54	96
42	57	55	90
43	111	56	81
44	194	57	73
45	256	58	64
46	296	59	61
47	378	60	59
48	250	61	50
49	185	62	42
50	171	63	29
51	150	64	18
52	110	65	15

a) ¿El supuesto de una distribución de Poisson parece apropiado como modelo de probabilidad para este proceso? Use $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

8-8 PRUEBAS DE TABLAS DE CONTINGENCIA

En muchas ocasiones, los n elementos de una muestra de una población pueden clasificarse con base en dos criterios diferentes. Entonces es de interés saber si los dos métodos de clasificación son estadísticamente independientes; por ejemplo, puede considerarse la población de ingenieros titulados, y tal vez quiera determinarse si el salario inicial es independiente de las disciplinas académicas. Suponga que el primer método de clasificación tiene r niveles y que el segundo tiene c niveles. Será O_{ij} la frecuencia observada del nivel i del primer método de clasificación y

el nivel j del segundo método de clasificación. Los datos aparecerían, en general, como en la tabla 8-3. A una tabla como ésta suele llamársele **tabla de contingencia $r \times c$**.

Nos interesa probar la hipótesis de que los métodos de clasificación de renglón y columna son independientes. Si se rechaza esta hipótesis, se concluye que hay cierta interacción entre los dos criterios de clasificación. Es difícil obtener procedimientos exactos de prueba, pero un estadístico de prueba aproximado es válido cuando n es grande. Sea p_{ij} la probabilidad de que un elemento seleccionado al azar esté en la celda ij, dado que las dos clasificaciones son independientes. Entonces $p_{ij} = u_i v_j$, donde u_i es la probabilidad de que un elemento seleccionado al azar esté en la clase del renglón i y v_j es la probabilidad de que un elemento seleccionado al azar esté en la clase de la columna j. Ahora bien, con el supuesto de independencia, los estimadores de u_i y v_j son

$$\hat{u}_i = \frac{1}{n}\sum_{j=1}^{c} O_{ij}$$

$$\hat{v}_j = \frac{1}{n}\sum_{i=1}^{r} O_{ij} \tag{8-82}$$

Por lo tanto, la frecuencia esperada de cada celda es

$$E_{ij} = n\hat{u}_i\hat{v}_j = \frac{1}{n}\sum_{j=1}^{c} O_{ij}\sum_{i=1}^{r} O_{ij} \tag{8-83}$$

Entonces, para n grande, el estadístico

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{8-84}$$

tiene una distribución ji-cuadrada aproximada con $(r-1)(c-1)$ grados de libertad si la hipótesis nula es verdadera. Por lo tanto, la hipótesis de independencia se rechazaría si el valor observado del estadístico de prueba χ_0^2 excediera $\chi^2_{\alpha,\,(r-1)(c-1)}$.

Tabla 8-3 Tabla de contingencia $r \times c$

		Columnas			
		1	2	...	c
Renglones	1	O_{11}	O_{12}	...	O_{1c}
	2	O_{21}	O_{22}	...	O_{2c}
	⋮	⋮	⋮	⋮	⋮
	r	O_{r1}	O_{r2}	...	O_{rc}

Tabla 8-4 Datos observados en el ejemplo 8-20

	Plan de pensión			
Clasificación laboral	1	2	3	Totales
Trabajadores asalariados	160	140	40	340
Trabajadores por hora	40	60	60	160
Totales	200	200	100	500

EJEMPLO 8-20

Una compañía tiene que elegir entre tres planes de pensiones. La administración quiere saber si la preferencia por los planes es independiente de la clasificación laboral y desea usar $\alpha = 0.05$. En la tabla 8-4 se muestran las opiniones de una muestra aleatoria de 500 empleados.

Para encontrar las frecuencias esperadas, primero debe calcularse $\hat{u}_1 = (340/500) = 0.68$, $\hat{u}_2 = (160/500) = 0.32$, $\hat{v}_1 = (200/500) = 0.40$, $\hat{v}_2 = (200/500) = 0.40$ y $\hat{v}_3 = (100/500) = 0.20$. Ahora pueden calcularse las frecuencias esperadas con la ecuación 8-83. Por ejemplo, el número esperado de trabajadores asalariados que prefieren el plan de pensión 1 es

$$E_{11} = n\hat{u}_1\hat{v}_1 = 500(0.68)(0.40) = 136$$

En la tabla 8-5 se muestran las frecuencias esperadas.

Ahora puede aplicarse el procedimiento de prueba de hipótesis de ocho pasos en este problema.

1. La variable de interés es la preferencia de los empleados entre los planes de pensiones.
2. H_0: La preferencia es independiente de la clasificación laboral asalariado o por horas.
3. H_1: La preferencia no es independiente de la clasificación laboral asalariado o por horas.
4. $\alpha = 0.05$
5. El estadístico de prueba es

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

6. Puesto que $r = 2$ y $c = 3$, los grados de libertad de ji-cuadrada son $(r-1)(c-1) = (1)(2) = 2$, y se rechazaría H_0 si $\chi_0^2 > \chi_{0.05,2}^2 = 5.99$.

Tabla 8-5 Frecuencias esperadas en el ejemplo 8-20

	Plan de pensión			
Clasificación laboral	1	2	3	Totales
Trabajadores asalariados	136	136	68	340
Trabajadores por hora	64	64	32	160
Totales	200	200	100	500

7. Cálculos:

$$\chi_0^2 = \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= \frac{(160-136)^2}{136} + \frac{(140-136)^2}{136} + \frac{(40-68)^2}{68} + \frac{(40-64)^2}{64}$$

$$+ \frac{(60-64)^2}{64} + \frac{(60-32)^2}{32} = 49.63$$

8. Conclusiones: Puesto que $\chi_0^2 = 49.63 > \chi_{0.05,2}^2 = 5.99$, se rechaza la hipótesis de independencia y se concluye que la preferencia por los planes de pensiones no es independiente de la clasificación laboral. El valor P para $\chi_0^2 = 49.63$ es $P = 1.671 \times 10^{-11}$. (Este valor se calculó con una calculadora HP-48.) Se necesitarían análisis adicionales para explorar la naturaleza de la asociación entre estos factores. Podría resultar útil examinar la tabla de las frecuencias observadas menos las esperadas.

El uso de una tabla de contingencia para probar la independencia entre dos variables de clasificación en una muestra de una sola población de interés es tan sólo una de las aplicaciones de los métodos de tablas de contingencia. Otra situación común ocurre cuando hay r poblaciones de interés y cada población se divide en las mismas c categorías. Entonces se toma una muestra de la población i, y los conteos se anotan en las columnas apropiadas del renglón i. En esta situación quiere investigarse si las proporciones en las c categorías son iguales o no para todas las poblaciones. La hipótesis nula en este problema plantea que las poblaciones son **homogéneas** con respecto a las categorías. Por ejemplo, cuando sólo hay dos categorías, como éxito y fracaso, defectuoso o no defectuoso, etc., la prueba de homogeneidad es en realidad una prueba de la igualdad de r parámetros binomiales. El cálculo de las frecuencias esperadas, la determinación de los grados de libertad, y el cálculo del estadístico ji-cuadrada para la prueba de homogeneidad son idénticos a la prueba de independencia.

EJERCICIOS DE LA SECCIÓN 8-8

8-64. Una compañía opera cuatro máquinas tres turnos diarios. Se recaban los siguientes datos de los registros de producción acerca del número de averías:

	Máquinas			
Turno	*A*	*B*	*C*	*D*
1	41	20	12	16
2	31	11	9	14
3	15	17	16	10

Pruebe la hipótesis (usando $\alpha = 0.05$) de que las averías son independientes del turno. Encuentre el valor P para esta prueba.

8-65. Los pacientes de un hospital se clasifican como quirúrgicos o médicos. Se lleva un registro del número de veces que los pacientes requieren atención médica durante la noche y si tienen seguro social. Los datos se presentan aquí:

Seguro social	Categoría del paciente	
	Quirúrgico	Médico
Sí	46	52
No	36	43

Pruebe la hipótesis (usando $\alpha = 0.01$) de que las consultas de los pacientes quirúrgicos-médicos son independiente de si los pacientes tienen seguro social. Encuentre el valor P para esta prueba.

8-66. Las calificaciones de un curso de estadística y de un curso de investigación de operaciones que se toman simultáneamente fueron las siguientes para un grupo de estudiantes:

Nota de estadística	Nota de investigación de operaciones			
	A	*B*	*C*	Otra
A	25	6	17	13
B	17	16	15	6
C	18	4	18	10
Otra	10	8	11	20

¿Las calificaciones de estadística e investigación de operaciones están relacionadas? Use $\alpha = 0.01$ para llegar a una conclusión. ¿Cuál es el valor P para esta prueba?

8-67. En un experimento con proyectiles de artillería se obtienen los siguientes datos sobre las características de desviación lateral y alcance. ¿Podría concluirse que la desviación y el alcance son independientes? Use $\alpha = 0.05$. ¿Cuál es el valor P de esta prueba?

Alcance (yardas)	Desviación lateral		
	Izquierda	Normal	Derecha
0-1 999	6	14	8
2 000-5 999	9	11	4
6 000-11 999	8	17	6

8-68. Se está haciendo un estudio de las fallas de un componente electrónico. Hay cuatro tipos de fallas posibles y dos posiciones de montaje para el dispositivo. Se han recabado los siguientes datos:

Posición de montaje	Tipo de falla			
	A	*B*	*C*	*D*
1	22	46	46	9
2	4	17	6	12

¿Podría concluirse que el tipo de falla es independiente de la posición de montaje? Use $\alpha = 0.01$. Encuentre el valor P para esta prueba.

8-69. Se pide a una muestra aleatoria de estudiantes universitarios su opinión sobre un cambio propuesto en el plan de estudios común. Los resultados son los siguientes:

Grado	Opinión	
	En favor	En contra
Primero	120	80
Segundo	70	130
Tercero	60	70
Cuarto	40	60

Pruebe la hipótesis de que la opinión acerca del cambio es independiente del año que cursan los estudiantes. Use $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

Ejercicios complementarios

8-70. Considere el intervalo de confianza para μ con desviación estándar σ conocida:

$$\hat{x} - z_{\alpha_1}\sigma/\sqrt{n} \le \mu \le \bar{x} + z_{\alpha_2}\sigma/\sqrt{n}$$

donde $\alpha_1 + \alpha_2 = \alpha$. Sea $\alpha = 0.05$ y encuentre el intervalo para $\alpha_1 = \alpha_2 = \alpha/2 = 0.025$. Encuentre ahora el intervalo para el caso en que $\alpha_1 = 0.01$ y $\alpha_2 = 0.04$. ¿Cuál de los dos intervalos es más corto? ¿Presenta alguna ventaja un intervalo de confianza "simétrico"?

8-71. Una población normal tiene una media conocida de 50 y una varianza desconocida.

a) Se selecciona una muestra aleatoria de $n = 16$ de esta población, y los resultados muestrales son $\bar{x} = 52$ y $s = 8$. ¿Qué tan extraños son estos resultados? Es decir, ¿cuál es la probabilidad de observar un

promedio muestral hasta de 52 (o mayor), si la media fundamental conocida es en realidad 50?

b) Se selecciona una muestra aleatoria de $n = 30$ de esta población, y los resultados muestrales son $\bar{x} = 52$ y $s = 8$. ¿Qué tan extraños son estos resultados?

c) Se selecciona una muestra aleatoria de $n = 100$ de esta población, y los resultados muestrales son $\bar{x} = 52$ y $s = 8$. ¿Qué tan extraños son estos resultados?

d) Compare las respuestas de los incisos a) al c) y explique por qué son iguales o por qué difieren.

8-72. Una población normal tiene una media $\mu = 50$ y una varianza $\sigma^2 = 5$ conocidas. ¿Cuál es la probabilidad aproximada de que la varianza muestral sea mayor o igual que 7.44? ¿Menor que 2.56?

a) Para una muestra aleatoria de $n = 16$.

b) Para una muestra aleatoria de $n = 30$.

c) Para una muestra aleatoria de $n = 71$.

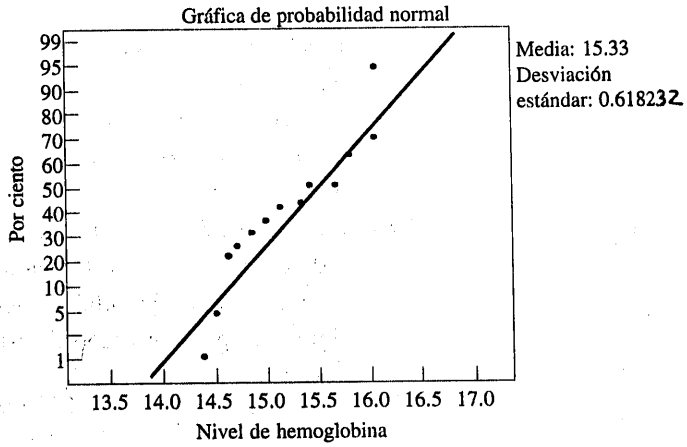
d) Compare las respuestas de los incisos a) al c) para la probabilidad aproximada de que la varianza muestral sea mayor o igual que 7.44. Explique por qué esta probabilidad de la cola es creciente o decreciente cuando se incrementa el tamaño de la muestra.

e) Compare las respuestas de los incisos a) al c) para la probabilidad aproximada de que la varianza muestral sea menor o igual que 2.56. Explique por qué esta probabilidad de la cola es creciente o decreciente cuando se incrementa el tamaño de la muestra.

8-73. Un artículo de *Journal of Sports Science* (Vol. 5, pp. 261-271) presenta los resultados de una investigación del nivel de hemoglobina de los jugadores del equipo olímpico de hockey de Canadá. Los datos reportados son los siguientes (en g/dl):

15.3	16.0	14.4	16.2	16.2
14.9	15.7	15.3	14.6	15.7
16.0	15.0	15.7	16.2	14.7
14.8	14.6	15.6	14.5	15.2

a) Dada la siguiente gráfica de probabilidad de los datos, ¿cuál es un supuesto lógico acerca de la distribución fundamental de los datos?



b) Explique por qué esta verificación de la distribución fundamental de los datos muestrales es importante si quiere construirse un intervalo de confianza para la media.

c) Con base en estos datos muestrales, un intervalo de confianza de 95% para la media es [15.04, 15.62]. ¿Es razonable inferir que la verdadera media podría ser 14.5? Explique su respuesta.

d) Explique por qué esta verificación de la distribución fundamental de los datos muestrales es importante si quiere construirse un intervalo de confianza para la varianza.

e) Con base en estos datos muestrales, un intervalo de confianza de 95% para la varianza es [0.22, 0.82]. ¿Es razonable inferir que la verdadera varianza podría ser 0.35? Explique su respuesta.

f) ¿Es razonable usar estos intervalos de confianza para hacer una inferencia acerca de la media y la varianza del nivel de hemoglobina

i) de los doctores canadienses? Explique su respuesta.

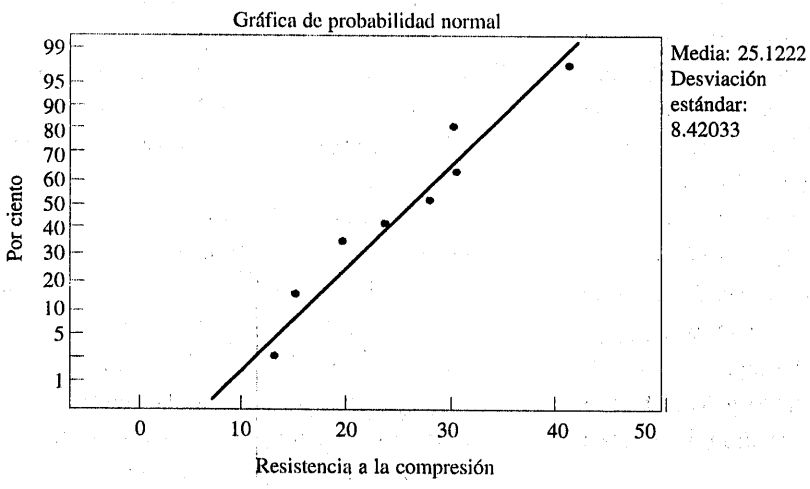
ii) de los niños canadienses entre 6 y 12 años de edad? Explique su respuesta.

8-74. En el artículo "Diseño mixto para el desarrollo de la fuerza óptima del concreto con ce-

nizas finas" (*Cement and Concrete Research*, Vol. 19, No. 4, pp. 634-640), se investiga la resistencia a la compresión del concreto cuando se mezcla con cenizas finas (una mezcla de sílice, alúmina, hierro, óxido de magnesio y otros ingredientes). La resistencia a la compresión de nueve muestras en condiciones secas en el día 28 son las siguientes (en MPa):

40.2	30.4	28.9	30.5	22.4
25.8	18.4	14.2	15.3	

a) Dada la siguiente gráfica de probabilidad de los datos, ¿cuál es un supuesto lógico acerca de la distribución fundamental de los datos?



b) Encuentre un intervalo de confianza inferior de 99% para la media de la resistencia a la compresión. Proponga una interpretación práctica de este intervalo.

c) Encuentre un intervalo de confianza de dos colas de 98% para la media de la resistencia a la compresión. Proponga una interpretación práctica de este intervalo y explique por qué el punto del extremo inferior del intervalo es el mismo o no que en el inciso b).

d) Encuentre un intervalo de confianza superior de 99% para la varianza de la resistencia a la compresión. Proponga una interpretación práctica de este intervalo.

e) Encuentre un intervalo de confianza de dos colas de 98% para la varianza de la resistencia a la compresión. Proponga una interpretación práctica de este intervalo y explique por qué el punto extremo superior del intervalo es el mismo o no que en el inciso d).

f) Suponga que se descubrió que la observación más grande 40.2 se registró incorrectamente y que es en realidad 20.4. Ahora la media muestral es $x = 23$ y la varianza muestral es $s^2 = 36.9$. Use estos nuevos valores y repita los incisos c) y e) anteriores. Compare los intervalos originales calculados y los intervalos recién calculados con el valor corregido de la observación. ¿Cómo afecta este error los valores de la media muestral, la varianza muestral y el ancho del intervalo de confianza de dos colas?

g) Suponga ahora que se descubrió que la observación más grande 40.2 es correcta, pero que la observación 25.8 es incorrecta y que es en realidad 24.8. Ahora la media muestral es $x = 25$ y la varianza muestral es $s^2 = 8.41$. Use estos nuevos valores y repita los incisos c) y e) anteriores. Compare los intervalos originales calculados y los intervalos recién calculados con el valor corregido de la observación. ¿Cómo afecta este error los valores de la media muestral, la varianza muestral y el ancho del intervalo de confianza de dos colas?

h) Use los resultados de los incisos f) y g) anteriores para explicar el efecto de los valores registrados incorrectamente sobre las estimaciones muestrales. Comente el efecto cuando los valores incorrectos están cerca de la media muestral y cuando no lo están.

8-75. El sistema operativo de una computadora personal se ha estudiado exhaustivamente, y se sabe que la desviación estándar del tiempo de respuesta después de un comando particular es $\sigma = 8$ milisegundos. Se instala una nueva versión del sistema operativo, y quiere estimarse el tiempo de respuesta medio del nuevo sistema para asegurar que un interva-

lo de confianza de 95% para μ tiene una longitud de a lo sumo 5 milisegundos.

a) Si puede suponerse que el tiempo de respuesta tiene una distribución normal y que $\sigma = 8$ para el nuevo sistema, ¿qué tamaño de la muestra recomendaría usted?

b) Suponga que el vendedor informa que la desviación estándar del tiempo de respuesta del nuevo sistema es menor, digamos $\sigma = 6$; dé el tamaño de la muestra que usted recomendaría y comente el efecto que tiene la desviación estándar menor sobre este cálculo.

c) Suponga que no es posible dar por hecho que el tiempo de respuesta del nuevo sistema esté distribuido normalmente, pero piense que puede seguir una distribución de Weibull. ¿Cuál es el tamaño mínimo de muestra que recomendaría para construir cualquier intervalo de confianza sobre la verdadera media del tiempo de respuesta?

8-76. Un fabricante de dispositivos semiconductores toma una muestra aleatoria de chips de tamaño n y los prueba, clasificando cada uno como defectuoso o no defectuoso. Sea $X_i = 0$ si el chip no está defectuoso y $X_i = 1$ si el chip está defectuoso. La fracción defectuosa de la muestra es

$$\hat{p}_i = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

¿Cuáles son la distribución de muestreo y las estimaciones de la media muestral y la varianza muestral de $\hat{p}$ cuando

a) el tamaño de la muestra es $n = 50$?

b) el tamaño de la muestra es $n = 80$?

c) el tamaño de la muestra es $n = 100$?

d) Compare las respuestas de los incisos a) al c) y comente el efecto del tamaño de la muestra sobre la varianza de la distribución de muestreo.

8-77. Considere la descripción del ejercicio 8-76. Después de sacar una muestra, el interés se encuentra en calcular el error al estimar el verdadero valor de p. Para cada tamaño de la muestra y cada estimación de p, calcule el error en el nivel de confianza de 95%.

a) $n = 50$ y $\hat{p} = 0.10$

b) $n = 80$ y $\hat{p} = 0.10$

c) $n = 100$ y $\hat{p} = 0.10$

d) Compare los resultados de los incisos a) al c) y comente el efecto del tamaño de la muestra sobre el error al estimar el verdadero valor de p y el nivel de confianza de 95%.

e) Repita los incisos a) al d), esta vez utilizando un nivel de confianza de 99%.

f) Examine los resultados cuando se usa el nivel de confianza de 95% y cuando se usa el nivel de confianza de 99% para calcular el error y explique lo que ocurre con la magnitud del error cuando se aumenta el porcentaje de confianza.

8-78. Un inspector de control de calidad de los dispositivos de medición de flujo usados para administrar fluidos intravenosos realizará una prueba de hipótesis para determinar si la rapidez media de flujo es diferente del ajuste de la velocidad de flujo de 200 ml/h. Con base en información anterior, la desviación estándar de la velocidad de flujo se supone conocida e igual a 12 ml/h. Para cada uno de los siguientes tamaños de la muestra, y $\alpha = 0.05$ fija, encuentre la probabilidad del error tipo II si la verdadera media es 205 ml/h.

a) $n = 20$

b) $n = 50$

c) $n = 100$

d) ¿La probabilidad del error tipo II aumenta o disminuye cuando se incrementa el tamaño de la muestra? Explique su respuesta.

8-79. En el ejercicio 8-78, suponga que el experimentador pensaba que $\sigma = 14$. Para cada uno de los siguientes tamaños de la muestra y $\alpha = 0.05$ fija, encuentre la probabilidad del error tipo II si la verdadera media es 205 ml/h.

a) $n = 20$

b) $n = 50$

c) $n = 100$

d) Al comparar las respuestas con las del ejercicio 8-74, ¿la probabilidad de un error tipo II aumenta o disminuye con el incremen-

to en la desviación estándar? Explique su respuesta.

8-80. Los comerciantes de champú saben que a los clientes les gusta que el producto haga mucha espuma. Un fabricante de champú afirma que la altura de la espuma de su producto excede 200 mm. Se sabe por experiencias anteriores que la desviación estándar de la altura de la espuma es 8 mm. Para cada uno de los siguientes tamaños de la muestra, y $\alpha = 0.05$ fija, encuentre la potencia de la prueba si la verdadera media es 204 mm.

a) $n = 20$
b) $n = 50$
c) $n = 100$
d) ¿La potencia de la prueba aumenta o disminuye cuando el tamaño de la muestra se incrementa? Explique su respuesta.

8-81. Suponga que quiere probarse la hipótesis H_0: $\mu = 85$ contra la hipótesis alternativa H_1: $\mu > 85$, donde $\sigma = 16$. Suponga que la verdadera media es $\mu = 86$ y que en el contexto práctico del problema ésta no es una desviación de $\mu_0 = 85$ que tenga significación práctica.

a) Para una prueba con $\alpha = 0.01$, calcule β para los tamaños de la muestra $n = 25$, 100, 400 y 2 500, suponiendo que $\mu = 86$.
b) Suponga que el promedio muestral es $\bar{x} = 86$. Encuentre el valor P para el estadístico de prueba para los diferentes tamaños de la muestra especificados en el inciso a). ¿Los datos serían estadísticamente significativos con $\alpha = 0.01$?
c) Comente el uso de un tamaño de la muestra grande en este problema.

8-82. El sistema de enfriamiento de un submarino nuclear se compone de un montaje de tubos soldados a través de los cuales se hace circular un agente refrigerante. Las especificaciones requieren que la resistencia de la soldadura debe alcanzar o exceder 150 psi.

a) Suponga que los ingenieros de diseño deciden probar la hipótesis H_0: $\mu = 150$ contra H_1: $\mu > 150$. Explique por qué esta elección de la hipótesis alternativa es mejor que H_1: $\mu < 150$.
b) Una muestra aleatoria de 20 soldaduras da como resultado $\bar{x} = 153.7$ psi y $s = 11.3$ psi. ¿Qué conclusiones pueden sacarse acerca de las hipótesis del inciso a)? Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

8-83. Suponga que se está probando H_0: $p = 0.5$ contra H_1: $p \neq 0.5$. Suponga que p es el verdadero valor de la proporción de la población.

a) Usando $\alpha = 0.05$, encuentre la potencia de la prueba para $n = 100$, 150 y 300 suponiendo que $p = 0.6$. Comente el efecto del tamaño de la muestra sobre la potencia de la prueba.
b) Usando $\alpha = 0.01$, encuentre la potencia de la prueba para $n = 100$, 150 y 300 suponiendo que $p = 0.6$. Compare las respuestas con las del inciso a) y comente el efecto de α sobre la potencia de la prueba para diferentes tamaños de la muestra.
c) Usando $\alpha = 0.05$, encuentre la potencia de la prueba para $n = 100$, suponiendo $p = 0.08$. Compare la respuesta con la del inciso a) y comente el efecto de la potencia de la prueba para el mismo tamaño de la muestra y nivel.
d) Usando $\alpha = 0.01$, ¿qué tamaño de la muestra se necesita si $p = 0.6$ y se quiere que $\beta = 0.05$? ¿Qué tamaño de la muestra se necesita si $p = 0.8$ y se quiere que $\beta = 0.05$? Compare los dos tamaños de la muestra y comente el efecto del verdadero valor de p sobre el tamaño de la muestra requerido cuando β se mantiene aproximadamente constante.

8-84. Considérese la brillantez del cinescopio del experimento descrito en el ejercicio 8-30.

a) Para el tamaño de la muestra $n = 10$, ¿los datos apoyan la afirmación de que la desviación estándar de la corriente es menor que 20 microamperes?
b) Suponga que en lugar de $n = 10$, el tamaño de la muestra fue 51. Repita el análisis realizado en el inciso a) usando $n = 51$.

c) Compare las respuestas y comente cómo afecta el tamaño de la muestra las conclusiones sacadas en los incisos a) y b).

8-85. Considérense las mediciones del ácido graso de la margarina dietética que se describen en el ejercicio 8-31.

a) Para el tamaño de la muestra $n = 6$, y usando una hipótesis alternativa de dos colas y $\alpha = 0.01$, pruebe H_0: $\sigma^2 = 1.0$.

b) Suponga que en lugar de $n = 6$, el tamaño de la muestra fue $n = 51$. Repita el análisis realizado en el inciso a) usando $n = 51$.

c) Compare las respuestas y comente cómo el tamaño de la muestra afecta las conclusiones obtenidas en los incisos a) y b).

8-86. Un fabricante de instrumentos de medición de precisión afirma que la desviación estándar en el uso de los instrumentos es a lo sumo 0.00002 mm. Un analista, que no está enterado de la afirmación anterior, usa el instrumento ocho veces y obtiene una desviación estándar muestral de 0.00001 mm.

a) Confirme mediante un procedimiento de prueba y un nivel α de 0.01 que no hay evidencia suficiente en favor de la afirmación de que la desviación estándar es a lo sumo 0.00002. Enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

b) Explique por qué la desviación estándar muestral, $s = 0.00001$, es menor que 0.00002, no obstante que los resultados del procedimiento de prueba estadística no apoyan la afirmación.

8-87. Una compañía de biotecnología produce una droga terapéutica cuya concentración tiene una desviación estándar de 4 g/l. Se ha propuesto un nuevo método para producir esta droga, aun cuando implica ciertos costos adicionales. La administración autorizará el cambio de la técnica de producción sólo si la desviación estándar de la concentración en el proceso nuevo es menor que 4 g/l. Los investigadores escogen $n = 10$ y obtienen los siguientes datos. Realice el análisis necesario para determinar si deberá implementarse el cambio en la técnica de producción.

16.628 g/l	16.630 g/l
16.622	16.631
16.627	16.624
16.623	16.622
16.618	16.626

8-88. Un fabricante de calculadoras electrónicas afirma que menos del 1% de su producción está defectuosa. Una muestra aleatoria de 1 200 calculadoras contiene ocho unidades defectuosas.

a) Confirme mediante un procedimiento de prueba y un nivel α de 0.01 que no hay evidencia suficiente para apoyar la afirmación de que el porcentaje de unidades defectuosas es menor que 1%.

b) Explique por qué el porcentaje muestral es menor que 1%, no obstante que los resultados del procedimiento de prueba no apoyan la afirmación.

8-89. Un artículo de *The Engineer* ("Rediseño de tendido eléctrico dudoso") reportó los resultados de una investigación de los errores de tendido eléctrico en el transporte aéreo comercial que pueden producir información incorrecta para la tripulación. Quizás uno de estos errores de tendido eléctrico haya ocasionado el accidente de un avión de British Midland Airways en enero de 1989, al hacer que el piloto apagara el motor equivocado. De 1 600 aviones seleccionados al azar, se encontró que ocho presentaban errores de tendido eléctrico que podían presentar información incorrecta a la tripulación.

a) Encuentre un intervalo de confianza de 99% para la proporción de aviones que presentan estos errores de tendido eléctrico.

b) Suponga que se usa la información de este ejemplo para ofrecer una estimación preliminar de p. ¿Cuál sería el tamaño necesario de una muestra para producir una estimación de p de la que se tuviera una con-

fianza de 99% de que difiere del valor real en a lo sumo 0.008?

c) Suponga que no se cuenta con una estimación preliminar de p. ¿Cuál sería el tamaño necesario de una muestra si se quisiera tener una confianza de al menos 99% de que la proporción muestral difiere de la proporción real en a lo sumo 0.008, independientemente del verdadero valor de p?

d) Comente la utilidad de la información preliminar para calcular el tamaño de la muestra que se necesita.

8-90. Un artículo de *Engineering Horizons* informó que 117 de 484 ingenieros titulados tenían contemplado continuar estudios de posgrado. Considere que esta información constituye una muestra aleatoria de la generación de graduados de 1990.

a) Encuentre un intervalo de confianza de 90% para la proporción de los graduados que planean continuar sus estudios.

b) Encuentre un intervalo de confianza de 95% para la proporción de los graduados que planean continuar sus estudios.

c) Compare las respuestas de los incisos a) y b) y explique por qué son iguales o diferentes.

d) ¿Podría usarse alguno de estos dos intervalos de confianza para determinar si la proporción difiere de manera significativa de 0.25? Explique su respuesta. [*Sugerencia:* Use la aproximación normal de la binomial.]

8-91. Considere las 40 observaciones recabadas del número de resortes en espiral que no cumplen con las especificaciones en lotes de producción de tamaño 50 dadas en el ejercicio 2-33 del capítulo 2.

a) Con base en la descripción de la variable aleatoria y en estas 40 observaciones, ¿una distribución binomial es un modelo apropiado? Realice un procedimiento de bondad del ajuste con $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

8-92. Considere las 20 observaciones registradas sobre el número de errores en una cadena de 1 000 bits de un canal de comunicación dadas en el ejercicio 2-34 del capítulo 2.

a) Con base en la descripción de la variable aleatoria y en estas 20 observaciones, ¿una distribución binomial es un modelo apropiado? Realice un procedimiento de bondad del ajuste con $\alpha = 0.05$.

b) Calcule el valor P para esta prueba.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

8-93. Un componente eléctrico tiene una distribución del tiempo hasta una falla (o vida útil) que es exponencial con parámetro λ, de tal modo que la vida útil media es $\mu = 1/\lambda$. Suponga que una muestra de n de estos componentes se somete a prueba, y sea X_i la vida útil observada del componente i. La prueba se continúa sólo hasta que la unidad r falla, donde $r < n$. Se obtiene así una prueba de la vida útil **censurada**. Sea que X_1 denote el tiempo en que ocurre la primera falla, que X_2 denote el tiempo en que ocurre la segunda falla, y así sucesivamente. Entonces la vida útil total que se ha acumulado al concluir la prueba es

$$T_r = \sum_{i=1}^{r} X_i + (n-r)X_r$$

a) Demuestre que T_r/r es un estimador insesgado de μ. [*Sugerencia:* Use la propiedad de falta de memoria de la distribución exponencial y el hecho de que si X_i ($i = 1, 2, ..., k$) son variables aleatorias exponenciales independientes con parámetro λ,

EJERCICIOS PARA DESARROLLAR EL INTELECTO

mín(X_1, X_2, ..., X_k) es exponencial con parámetro $k\lambda$.]

b) Puede demostrarse que $2\lambda T_r$ tiene una distribución ji-cuadrada con $2r$ grados de libertad. Use este hecho para desarrollar un intervalo de confianza del $100(1 - \alpha)$ por ciento para la vida útil media $\mu = 1/\lambda$.

c) Suponga que se sometieron a prueba 20 unidades y que la prueba terminó después de que ocurrieron 10 fallas. Los tiempos de falla (en horas) son: 15, 18, 19, 20, 21, 21, 22, 27, 28, 29. Encuentre un intervalo de confianza de 95% para la vida útil media.

8-94. Considere un intervalo de confianza de dos colas para la media μ cuando σ es conocida:

$$\bar{x} - z_{\alpha_1}\sigma/\sqrt{n} \le \mu \le \bar{x} + z_{\alpha_2}\sigma/\sqrt{n}$$

donde $\alpha_1 + \alpha_2 = \alpha$. Si $\alpha_1 = \alpha_2 = \alpha/2$, se tiene el intervalo de confianza del $100(1 - \alpha)$ por ciento usual para μ. En lo anterior, cuando $\alpha_1 \neq \alpha_2$, el intervalo no es simétrico alrededor de μ. La longitud del intervalo es $L = \sigma(z_{\alpha_1} + z_{\alpha_2})/\sqrt{n}$. Demuestre que la longitud del intervalo L se minimiza cuando $\alpha_1 = \alpha_2 = \alpha/2$. [*Sugerencia:* recuerde que $\Phi(z_\alpha) = 1 - \alpha$, por lo que $\Phi^{-1}(1 - \alpha) = z_\alpha$, y que la relación entre la derivada de una función $y = f(x)$ y la inversa $x = f^{-1}(y)$ es $(d/dy)\, f^{-1}(y) = 1/[(d/dx)f(x)]$.]

8-95. Intervalo de predicción para una observación futura. Este capítulo se ha enfocado en los intervalos de confianza para un solo parámetro, como la media. Sin embargo, en muchas situaciones prácticas querrá predecirse un solo valor futuro de la variable aleatoria. En este problema se pide que se desarrolle el estimador del intervalo apropiado para esta predicción.

Considere nuevamente la situación del ejercicio 8-33, donde se seleccionó una muestra aleatoria de 15 varillas para la suspensión de un automóvil y se midieron sus diámetros. En ese ejercicio se pedía encontrar un intervalo de confianza de 95% para el diámetro medio. Suponga sin embargo, que usted instalará una varilla de este tipo en su automóvil y que quiere una predicción del diámetro de esa varilla específica.

a) Explique por qué el promedio muestral $\bar{X}$, basado en la muestra aleatoria original de tamaño n, es un predictor puntual razonable de la observación futura, digamos X_{n+1}.

b) Si el predictor puntual es $\bar{X}$, entonces el error de predicción es $X_{n+1} - \bar{X}$. Demuestre que el error de predicción esperado es cero y que la varianza del error de predicción es $\sigma^2\left(1 + \frac{1}{n}\right)$.

c) Si S es la desviación estándar muestral basada en las n observaciones originales, determine la distribución de

$$\frac{X_{n+1} - \bar{X}}{S\sqrt{1 + \frac{1}{n}}}$$

d) Use el resultado del inciso c) para demostrar que un intervalo de predicción del $100(1 - \alpha)$ por ciento para la observación futura única es

$$\bar{x} - t_{\alpha/2,\, n-1}\, s\sqrt{1 + \frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{\alpha/2,\, n-1}\, s\sqrt{1 + \frac{1}{n}}$$

e) Use los datos de los diámetros de las varillas para calcular el intervalo de predicción de 95% para el diámetro de una varilla futura.

f) Compare el intervalo obtenido en el inciso e) anterior con el intervalo de confianza para la media calculado en el ejercicio

8-33. Obsérvese que el intervalo de predicción tiene una anchura mayor. ¿Por qué? Explique la diferencia al interpretar estos dos intervalos.

8-96. Intervalo de tolerancia. Suponga que X tiene una distribución normal con media μ y varianza σ^2. Entonces el intervalo $\mu \pm 1.645\sigma$ incluirá 95% de los valores de X. A este tipo de intervalo se le llama intervalo de tolerancia. En la práctica, estos intervalos son muy útiles, ya que definen un rango en el que estará incluida una proporción establecida de la variable de interés. Desafortunadamente, por lo general no se conocen μ ni σ, por lo que el intervalo de tolerancia se calcula como $\overline{X} \pm kS$, donde k es una constante elegida de tal modo que se tenga una confianza de 90% (por ejemplo) de que el intervalo incluye 95% de los valores de la variable de interés. La constante k se tabula adecuadamente para varios valores del tamaño de la muestra n, el nivel de confianza y la probabilidad de cobertura (95%, en este caso). Por ejemplo, véase Montgomery.

a) Considere los datos del diámetro de las varillas para la suspensión de un automóvil del ejercicio 8-33. Para $n = 15$, una confianza de 90%, y una cobertura de 95%, $k = 2.713$. Calcule el intervalo de tolerancia que incluya al menos 95% de los diámetros de las varillas con una confianza de 90%.

b) Explique en qué forma difiere este intervalo del intervalo de confianza para la media del diámetro de las varillas.

c) ¿Cuál es el comportamiento de la longitud de un intervalo de tolerancia cuando n se hace grande? Obsérvese que para una confianza de 90% y una cobertura de 95%, $k = 2.172$ para $n = 100$ y $k = 2.019$ para $n = 1\,000$. ¿Qué valor de k sería adecuado cuando $n \to \infty$? ¿Cuál es el comportamiento de la longitud de un intervalo de tolerancia cuando $n \to \infty$?

8-97. Es posible construir un **intervalo de tolerancia no paramétrico** que se base en los valores extremos de una muestra aleatoria de tamaño n de cualquier población continua. Si p es la proporción mínima de la población contenida entre las observaciones muestrales menor y mayor con una confianza $1 - \alpha$, entonces puede demostrarse que

$$np^{n-1} - (n-1)p^n = \alpha$$

y n es aproximadamente

$$n = \frac{1}{2} + \left(\frac{1+p}{1-p}\right)\left(\frac{\chi^2_{\alpha,4}}{4}\right)$$

a) A fin de tener una confianza de 95% de que al menos 90% de la población estará incluida entre los valores extremos de la muestra, ¿qué tamaño de la muestra se necesitará?

b) Una muestra aleatoria de 10 transistores dio las siguientes mediciones de corriente de saturación (en miliamperes): 10.25, 10.41, 10.30, 10.26, 10.19, 10.37, 10.29, 10.34, 10.23, 10.38. Encuentre los límites que contienen una proporción p de las mediciones de la corriente de saturación con una confianza de 95%. ¿Cuál es la proporción p contenida en estos límites?

8-98. Suponga que $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución de probabilidad continua con mediana $\tilde{\mu}$.

a) Demuestre que

$$P\{\text{mín}\,(X_i) < \tilde{\mu} < \text{máx}\,(X_i)\} = 1 - \left(\frac{1}{2}\right)^{n-1}$$

[*Sugerencia:* El complemento del evento [mín(X_i) < $\tilde{\mu}$ < máx(X_i)] es [máx(X_i) $\leq \tilde{\mu}$] $\cup$ [mín(X_i) $\leq \tilde{\mu}$], pero máx(X_i) $\leq \mu$ si y sólo si $X_i \leq \tilde{\mu}$ para toda i.]

b) Construya un intervalo de confianza del $100(1 - \alpha)$ por ciento para la mediana μ,

EJERCICIOS PARA DESARROLLAR EL INTELECTO

donde $\alpha = \left(\frac{1}{2}\right)^{n-1}$.

8-99. Suponga que quiere probarse H_0: $\mu = \mu_0$ contra H_1: $\mu \neq \mu_0$, donde la población es normal con σ conocida. Sea $0 < \varepsilon < \alpha$, y la región crítica se define de tal modo que se rechazará H_0 si $z_0 > z_\varepsilon$ o si $z_0 < -z_{\alpha-\varepsilon}$, donde z_0 es el valor del estadístico de prueba acostumbrado para estas hipótesis.

a) Demuestre que la probabilidad del error tipo I para esta prueba es α.

b) Suponga que la verdadera media es $\mu_1 = \mu_0 + \delta$. Deduzca una expresión para β para la prueba anterior.

8-100. Deduzca una expresión para β para la prueba sobre la varianza de una distribución normal. Suponga que se especifica la hipótesis alternativa de dos colas.

9 Inferencia estadística para dos muestras

PLAN GENERAL DEL CAPÍTULO

9-1 INTRODUCCIÓN

9-2 INFERENCIA PARA LA DIFERENCIA EN MEDIAS, VARIANZAS CONOCIDAS

- 9-2.1 Pruebas de hipótesis para la diferencia en medias, varianzas conocidas
- 9-2.2 Elección del tamaño de la muestra
- 9-2.3 Identificación de causa y efecto
- 9-2.4 Intervalo de confianza para una diferencia en medias, varianzas conocidas

9-3 INFERENCIA PARA LA DIFERENCIA EN MEDIAS DE DOS DISTRIBUCIONES NORMALES, VARIANZAS DESCONOCIDAS

- 9-3.1 Pruebas de hipótesis para la diferencia en medias, varianzas desconocidas
- 9-3.2 Elección del tamaño de la muestra
- 9-3.3 Intervalo de confianza para la diferencia en medias
- 9-3.4 Solución de computadora

9-4 PRUEBA t PAREADA

9-5 INFERENCIA SOBRE LAS VARIANZAS DE DOS POBLACIONES NORMALES

- 9-5.1 Pruebas de hipótesis sobre el cociente de dos varianzas
- 9-5.2 Error β y elección del tamaño de la muestra
- 9-5.3 Intervalo de confianza sobre el cociente de dos varianzas

9-6 INFERENCIA SOBRE PROPORCIONES DE DOS POBLACIONES

- 9-6.1 Prueba de una muestra grande para $H_0 : p_1 = p_2$
- 9-6.2 Error β y elección del tamaño de la muestra
- 9-6.3 Intervalo de confianza para $p_1 - p_2$

9-7 TABLA DE RESUMEN PARA PROCEDIMIENTOS DE INFERENCIA PARA DOS MUESTRAS

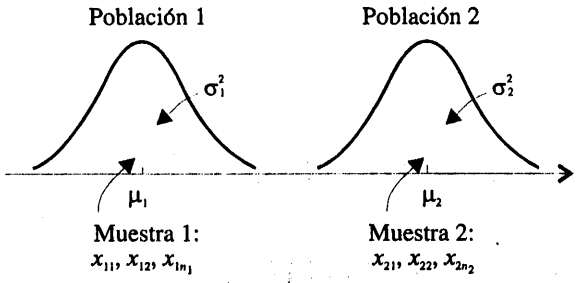


Figura 9-1 Dos poblaciones independientes.

9-1 INTRODUCCIÓN

En el capítulo anterior se presentaron las pruebas de hipótesis y los intervalos de confianza para un solo parámetro poblacional (la media μ, la varianza σ^2 o una proporción p). En este capítulo se amplían esos resultados al caso de dos poblaciones independientes.

En la figura 9-1 se ilustra la situación general. La población 1 tiene media μ_1 y varianza σ_1^2, mientras que la población 2 tiene media μ_2 y varianza σ_2^2. Las inferencias se basarán en dos muestras aleatorias de tamaños n_1 y n_2, respectivamente. Es decir, $X_{11}, X_{12}, \ldots, X_{1n_1}$ es una muestra aleatoria de n_1 observaciones de la población 1, y $X_{21}, X_{22}, \ldots, X_{1n_2}$ es una muestra aleatoria de n_2 observaciones de la población 2.

9-2 INFERENCIA PARA LA DIFERENCIA EN MEDIAS, VARIANZAS CONOCIDAS

En esta sección se consideran las inferencias estadísticas sobre la diferencia de las medias $\mu_1 - \mu_2$ de las poblaciones que se muestran en la figura 9-1, donde las varianzas σ_1^2 y σ_2^2 son conocidas. Los supuestos que se usarán en esta sección se resumen a continuación:

Supuestos

1. $X_{11}, X_{12}, \ldots, X_{1n_1}$ es una muestra aleatoria de la población 1.
2. $X_{21}, X_{22}, \ldots, X_{2n_2}$ es una muestra aleatoria de la población 2.
3. Las dos poblaciones representadas por X_1 y X_2 son independientes.
4. Ambas poblaciones son normales, o si no lo son, se cumplen las condiciones del teorema del límite central.

Un estimador puntual lógico de $\mu_1 - \mu_2$ es la diferencia de las medias muestrales $\overline{X}_1 - \overline{X}_2$. Con base en las propiedades de los valores esperados del capítulo 6, se tiene

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

y la varianza de $\bar{X}_1 - \bar{X}_2$ es

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Con base en los supuestos y los resultados precedentes, puede enunciarse lo siguiente:

La cantidad

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \tag{9-1}$$

tiene una distribución $N(0, 1)$.

Este resultado se usará para hacer pruebas de hipótesis y construir intervalos de confianza para $\mu_1 - \mu_2$. En esencia, $\mu_1 - \mu_2$ puede considerarse como un parámetro θ, cuyo estimador es $\hat{\Theta} = \bar{X}_1 - \bar{X}_2$ con varianza $\sigma_{\hat{\Theta}}^2 = \sigma_1^2/n_1 + \sigma_2^2/n_2$. Si θ_0 es el valor especificado para θ en la hipótesis nula, entonces el estadístico de la prueba será $(\hat{\Theta} - \theta_0)/\sigma_{\hat{\Theta}}$. Obsérvese la gran similitud entre este resultado y el estadístico de la prueba para una sola media usado en la ecuación 8-10 del capítulo 8.

9-2.1 Pruebas de hipótesis para la diferencia en medias, varianzas conocidas

Se considera ahora la prueba de hipótesis sobre la diferencia de las medias $\mu_1 - \mu_2$ de las dos poblaciones de la figura 9-1. Suponga que el interés está en probar que la diferencia de las medias $\mu_1 - \mu_2$ es igual a un valor especificado Δ_0. Por tanto, la hipótesis nula se enunciará como H_0: $\mu_1 - \mu_2 = \Delta_0$. Desde luego, en muchos casos se especificará $\Delta_0 = 0$, de tal modo que se estaría probando la igualdad de las dos medias (es decir, H_0: $\mu_1 = \mu_2$). El estadístico de la prueba apropiado se encontraría sustituyendo $\mu_1 - \mu_2$ con Δ_0 en la ecuación 9-1, y este estadístico de la prueba tendría una distribución normal estándar bajo H_0. Suponga que la hipótesis alternativa es H_1: $\mu_1 - \mu_2 \neq \Delta_0$. Ahora bien, un valor muestral de $\bar{x}_1 - \bar{x}_2$ que difiera considerablemente de Δ_0 es evidencia de que H_1 es verdadera. Debido a que Z_0 tiene la distribución $N(0, 1)$ cuando H_0 es verdadera, se tomarían $-z_{\alpha/2}$ y $z_{\alpha/2}$ como los límites de la región crítica justo como se hizo en el problema de la prueba de hipótesis para una sola muestra de la sección 8-2.1. Se obtendría así una prueba con nivel de significación α. Las regiones críticas para las hipótesis alternativas de una cola se localizarían de manera similar. A continuación se hace el resumen formal de estos resultados.

Prueba de hipótesis sobre $\mu_1 - \mu_2$, varianzas conocidas

Hipótesis nula: $H_0 : \mu_1 - \mu_2 = \Delta_0$

Estadístico de la prueba:
$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \qquad (9\text{-}2)$$

Hipótesis alternativas	Criterio de rechazo
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$z_0 > z_{\alpha/2}$ o $z_0 < -z_{\alpha/2}$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$z_0 > z_\alpha$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$z_0 < -z_\alpha$

EJEMPLO 9-1

El encargado de desarrollar un producto está interesado en reducir el tiempo de secado de una pintura tapaporo. Se prueban dos formulaciones de la pintura; la formulación 1 incluye las sustancias químicas comunes y la formulación 2 tiene un nuevo ingrediente que deberá reducir el tiempo de secado. Se sabe por experiencia que la desviación estándar del tiempo de secado es 8 minutos, y la adición del nuevo ingrediente no deberá afectar esta variabilidad inherente. Se pintan 10 muestras con la formulación 1, y otras 10 diez muestras se pintan con la formulación 2; las 20 muestras se pintan en orden aleatorio. Los dos promedios muestrales del tiempo de secado son $\bar{x}_1 = 121$ min y $\bar{x}_2 = 112$ min, respectivamente. ¿A qué conclusiones puede llegar el encargado del desarrollo del producto acerca de la efectividad del nuevo ingrediente usando $\alpha = 0.05$?

Se aplica el procedimiento de ocho pasos en este problema de la manera siguiente:

1. La cantidad de interés es la diferencia en las medias del tiempo de secado, $\mu_1 - \mu_2$, y $\Delta_0 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, o H_0: $\mu_1 = \mu_2$.
3. H_1: $\mu_1 > \mu_2$. Quiere rechazarse H_0 si el nuevo ingrediente reduce la media del tiempo de secado.
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$z_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

   donde $\sigma_1^2 = \sigma_2^2 = (8)^2 = 64$ y $n_1 = n_2 = 10$.
6. Se rechaza H_0: $\mu_1 = \mu_2$ si $z_0 > 1.645 = z_{0.05}$.

7. Cálculos: Puesto que $\bar{x}_1 = 121$ min y $\bar{x}_2 = 112$ min, el estadístico de la prueba es

$$z_0 = \frac{121-112}{\sqrt{\frac{(8)^2}{10}+\frac{(8)^2}{10}}} = 2.52$$

8. Conclusión: Puesto que $z_0 = 2.52 > 1.645$, se rechaza H_0: $\mu_1 = \mu_2$ con el nivel $\alpha = 0.05$, y se concluye que la adición del nuevo ingrediente a la pintura reduce significativamente el tiempo de secado. Alternativamente, puede encontrarse el valor P para esta prueba como

$$\text{Valor } P = 1 - \Phi(2.52) = 0.0059$$

Por lo tanto, H_0: $\mu_1 = \mu_2$ se rechazaría con cualquier nivel de significación $\alpha \geq 0.0059$.

9-2.2 Elección del tamaño de la muestra

Uso de las curvas de operación característica

Pueden usarse las curvas de operación característica de las gráficas VI*a*, VI*b*, VI*c* y VI*d* del apéndice para evaluar la probabilidad del error tipo II de las hipótesis del recuadro (9-2). Estas curvas también son útiles para determinar el tamaño de la muestra. Se incluyen curvas para $\alpha = 0.05$ y $\alpha = 0.01$. Para las hipótesis alternativas de dos colas, la escala de las abscisas de la curva de operación característica de las gráficas VI*a* y VI*b* es d, donde

$$d = \frac{|\mu_1 - \mu_2 - \Delta_0|}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{|\Delta - \Delta_0|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{9-3}$$

y debe elegirse el mismo tamaño de las muestras, es decir, $n = n_1 = n_2$. Las hipótesis alternativas de una cola requieren el uso de las gráficas VI*c* y VI*d*. Para la hipótesis alternativa de una cola H_1: $\mu_1 - \mu_2 > \Delta_0$ la escala de las abscisas es

$$d = \frac{\mu_1 - \mu_2 - \Delta_0}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{\Delta - \Delta_0}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{9-4}$$

con $n = n_1 = n_2$. La otra hipótesis alternativa de una cola, H_1: $\mu_1 - \mu_2 < \Delta$, requiere que d se defina como

$$d = \frac{\mu_2 - \mu_1 - \Delta_0}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{\Delta - \Delta_0}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{9-5}$$

y $n = n_1 = n_2$.

No es raro encontrar problemas donde los costos para recabar los datos difieren de manera sustancial entre las dos poblaciones, o donde la varianza de una población es mucho mayor que

la otra. En estos casos, suele usarse un tamaño diferente para cada muestra. Si $n_1 \neq n_2$, las curvas de operación característica pueden utilizarse con un valor *equivalente* de n que se calcula con

$$n = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2/n_1 + \sigma_2^2/n_2} \tag{9-6}$$

Si $n_1 \neq n_2$, y sus valores se fijan de antemano, la ecuación 9-6 se usa directamente para calcular n, y las curvas de operación característica se utilizan con una d especificada para obtener β. Si se da d y es necesario determinar n_1 y n_2 para obtener una β especificada, por ejemplo β^*, entonces se suponen valores de prueba de n_1 y n_2, se calcula n con la ecuación 9-6 y se utilizan las curvas con el valor especificado de d para encontrar β. Si $\beta = \beta^*$, entonces los valores supuestos de n_1 y n_2 son satisfactorios. Si $\beta \neq \beta^*$, entonces se hacen ajustes a n_1 y n_2 y se repite el proceso.

EJEMPLO 9-2

Considérese el experimento del tiempo de secado de la pintura del ejemplo 9-1. Si la verdadera diferencia en los tiempos promedio de secado es hasta de 10 minutos, encuentre los tamaños de las muestras requeridos para detectar esta diferencia con una probabilidad de al menos 0.90.

El valor apropiado del parámetro de las abscisas es (puesto que $\Delta_0 = 0$, y $\Delta = 10$)

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{10}{\sqrt{8^2 + 8^2}} = 0.88$$

y como la probabilidad de detección o potencia de la prueba debe ser al menos 0.9, con $\alpha = 0.05$, en la gráfica VI*c* del apéndice se encuentra que $n = n_1 = n_2 \simeq 11$.

Fórmulas del tamaño de la muestra

También es posible obtener fórmulas para calcular directamente el tamaño de las muestras. Suponga que la hipótesis nula H_0: $\mu_1 - \mu_2 = \Delta_0$ es falsa y que la verdadera diferencia de las medias es $\mu_1 - \mu_2 = \Delta$, donde $\Delta > \Delta_0$. Pueden encontrarse fórmulas del tamaño de la muestra que se necesita para obtener un valor específico de la probabilidad β del error tipo II para una diferencia de las medias Δ y un nivel de significación α dados.

Para las hipótesis alternativas de dos colas con nivel de significación α, el tamaño de la muestra $n_1 = n_2 = n$ que se necesita para detectar una diferencia real de las medias de Δ con una potencia de al menos $1 - \beta$ es

$$n \simeq \frac{(z_{\alpha/2} + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} \tag{9-7}$$

Esta aproximación es válida cuando el valor de $\Phi\left(-z_{\alpha/2}-(\Delta-\Delta_0)\sqrt{n}/\sqrt{\sigma_1^2+\sigma_2^2}\right)$ es pequeño en comparación con β.

Para una hipótesis alternativa de una cola con nivel de significación α, el tamaño de la muestra $n_1 = n_2 = n$ que se necesita para detectar una diferencia real de las medias de $\Delta(\neq \Delta_0)$, con una potencia de al menos $1-\beta$ es

$$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2+\sigma_2^2)}{(\Delta-\Delta_0)^2} \tag{9-8}$$

La deducción de las ecuaciones 9-7 y 9-8 es muy similar al caso de una sola muestra tratado en la sección 8-2.3. Por ejemplo, para obtener la ecuación 9-7, primero se escribe la expresión del error β de la hipótesis alternativa de dos colas, que es

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\Delta-\Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta-\Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\right)$$

donde Δ es la verdadera diferencia de las medias de interés. Después, siguiendo un procedimiento similar al que se usó para obtener la ecuación 8-22, puede obtenerse la expresión para β para el caso en que $n = n_1 = n_2$.

EJEMPLO 9-3

Para ilustrar el uso de estas ecuaciones del tamaño de las muestras, considérese la situación descrita en el ejemplo 9-1, y suponga que si la verdadera diferencia de los tiempos de secado es a lo más de 10 minutos, quiere detectarse este hecho con una probabilidad de al menos 0.90. Bajo la hipótesis nula, $\Delta_0 = 0$. Se tiene una hipótesis alternativa de una cola con $\Delta = 10$, $\alpha = 0.05$ (por lo que $z_\alpha = z_{0.05} = 1.645$), y ya que la potencia es 0.9, $\beta = 0.10$ (por lo que $z_\beta = z_{0.10} = 1.28$). Por lo tanto, el tamaño de la muestra requerido puede encontrarse con la ecuación 9-8 de la siguiente manera:

$$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2+\sigma_2^2)}{(\Delta-\Delta_0)^2} = \frac{(1.645+1.28)^2\left[(8)^2+(8)^2\right]}{(10-0)^2}$$
$$= 11$$

Este resultado concuerda exactamente con el obtenido usando las curvas OC.

9-2.3 Identificación de causa y efecto

Con frecuencia ingenieros y científicos tienen interés en comparar dos condiciones diferentes a fin de determinar si alguna de las dos condiciones produce un efecto significativo en la respuesta que se observa. En ocasiones a estas condiciones se les llama **tratamientos**. En el ejemplo 9-1 se ilustra una situación como ésta; los dos tratamientos diferentes son las dos formulaciones de la pintura, y la respuesta es el tiempo de secado. El objetivo del estudio es determinar si la nueva formulación tiene como resultado un efecto significativo: reducir el tiempo de secado. En esta situación, el responsable de desarrollar el producto (el experimentador) asignó aleatoriamente 10 muestras de prueba a una formulación y 10 muestras de prueba a la otra formulación. Después se aplicó la pintura a las muestras de prueba en orden aleatorio hasta que se pintaron las 20 muestras. Se trata de un ejemplo de un **experimento totalmente aleatorizado**.

Cuando se observa una significación estadística en un experimento aleatorizado, el experimentador puede confiar en la conclusión de que fue la diferencia en los tratamientos lo que dio como resultado la diferencia en la respuesta. Es decir, puede confiarse en que se ha encontrado una relación de causa y efecto.

En ocasiones los objetos que se van a usar en la comparación no se asignan aleatoriamente a los tratamientos. Por ejemplo, en un número de *Circulation* (una revista médica publicada por la American Heart Association) se reportó un estudio que relacionaba los niveles altos de hierro en el cuerpo con un riesgo mayor de un infarto cardiaco. En el estudio, realizado en Finlandia, se hizo el seguimiento de 1 931 hombres durante cinco años y demostró un efecto estadísticamente significativo de los niveles altos de hierro en la incidencia de infartos cardiacos. En este estudio, la comparación no se realizó seleccionando una muestra aleatoria de hombres y asignando después algunos a un tratamiento de "nivel bajo de hierro" y al resto a un tratamiento de "nivel alto de hierro". Los investigadores se limitaron a hacer el seguimiento de los sujetos en el tiempo. Recuérdese por el capítulo 1 que a este tipo de estudio se le llama **estudio observacional**.

En los estudios observacionales es difícil identificar una relación de causalidad, ya que la diferencia estadísticamente significativa observada en la respuesta entre los dos grupos puede haberse debido a otro factor (o grupo de factores) fundamental, que no se compensó aleatorizando el proceso, y no a los tratamientos. Por ejemplo, la diferencia en el riesgo de un infarto cardiaco podría atribuirse a la diferencia en los niveles de hierro, o a otros factores fundamentales que forman una explicación razonable de los resultados observados, tales como los niveles de colesterol o la hipertensión.

La dificultad para establecer una relación de causalidad a partir de estudios observacionales también puede verse en la controversia entre tabaquismo y salud. Numerosos estudios indican que la incidencia de cáncer pulmonar y otras enfermedades respiratorias es más alta entre los fumadores que entre los no fumadores. Sin embargo, establecer una relación de causa y efecto en este caso ha demostrado ser en extremo difícil. Muchos individuos decidieron fumar mucho antes de que se iniciaran los estudios de investigación, y muchos factores, además del tabaquismo, podrían intervenir en los casos de cáncer pulmonar.

9-2.4 Intervalo de confianza para una diferencia en medias, varianzas conocidas

El intervalo de confianza del $100(1 - \alpha)$ por ciento para la diferencia de dos medias $\mu_1 - \mu_2$ cuando las varianzas son conocidas puede encontrarse directamente a partir de los resultados ya

presentados en esta sección. Recuérdese que $X_{11}, X_{12}, \ldots, X_{1n_1}$ es una muestra aleatoria de n_1 observaciones de la primera población y que $X_{21}, X_{22}, \ldots, X_{2n_2}$ es una muestra aleatoria de n_2 observaciones de la segunda población. La diferencia de las medias muestrales $X_1 - X_2$ es un estimador puntual de $\mu_1 - \mu_2$, y

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

tiene una distribución normal estándar si las dos poblaciones son normales o es aproximadamente normal estándar si se cumplen las condiciones del teorema del límite central, respectivamente. Esto implica que

$$P\left(-z_{\alpha/2} \le Z \le z_{\alpha/2}\right) = 1 - \alpha$$

o bien

$$P\left(-z_{\alpha/2} \le \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \le z_{\alpha/2}\right) = 1 - \alpha$$

Esta expresión puede reordenarse como

$$P\left(\bar{X}_1 - \bar{X}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \bar{X}_1 - \bar{X}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) = 1 - \alpha$$

Por lo tanto, el intervalo de confianza del $100(1 - \alpha)$ por ciento para $\mu_1 - \mu_2$ se define como sigue:

Definición: Intervalo de confianza para la diferencia en dos medias, varianzas conocidas

Si $\bar{x}_1$ y $\bar{x}_2$ son las medias de las muestras aleatorias independientes de tamaños n_1 y n_2 de poblaciones con varianzas conocidas σ_1^2 y σ_2^2, respectivamente, entonces un intervalo de confianza del $100(1 - \alpha)$ por ciento para $\mu_1 - \mu_2$ es

$$\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (9\text{-}9)$$

donde $z_{\alpha/2}$ es el punto porcentual $\alpha/2$ superior de la distribución normal estándar.

El nivel de confianza $1 - \alpha$ es exacto cuando las poblaciones son normales. Para poblaciones no normales, el nivel de confianza es válido aproximadamente para tamaños grandes de las muestras.

EJEMPLO 9-4

Se realizaron pruebas de resistencia a la tensión en largueros de aluminio usados en la fabricación del ala de un avión de transporte comercial, de dos calidades diferentes. Por experiencias pasadas con el proceso de fabricación de largueros y el procedimiento de prueba, se suponen conocidas las desviaciones estándar de las resistencias a la tensión. En la tabla 9-1 se muestran los datos obtenidos. Si μ_1 y μ_2 denotan las verdaderas medias de la resistencia a la tensión para las dos calidades de los largueros, entonces puede encontrarse un intervalo de confianza de 90% para la diferencia de la resistencia media $\mu_1 - \mu_2$ como sigue:

$$l = \bar{x}_1 - \bar{x}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= 87.6 - 74.5 - 1.645\sqrt{\frac{(1.0)^2}{10} + \frac{(1.5)^2}{12}}$$

$$= 13.1 - 0.88$$

$$= 12.22 \text{ kg/mm}^2$$

$$u = \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= 87.6 - 74.5 + 1.645\sqrt{\frac{(1.0)^2}{10} + \frac{(1.5)^2}{12}}$$

$$= 13.1 + 0.88$$

$$= 13.98 \text{ kg/mm}^2$$

Por lo tanto, el intervalo de confianza de 90% para la diferencia de la resistencia a la tracción media es

$$12.22 \text{ kg/mm}^2 \leq \mu_1 - \mu_2 \leq 13.98 \text{ kg/mm}^2$$

Obsérvese que el intervalo de confianza no incluye a cero, lo cual implica que la resistencia media del aluminio de calidad 1 (μ_1) excede la resistencia media del aluminio de calidad 2 (μ_2). De hecho, puede afirmarse que se tiene una confianza de 90% de que la resistencia a la tensión media del aluminio de calidad 1 excede la del aluminio calidad 2, por una cantidad entre 12.22 y 13.98 kg/mm^2.

Tabla 9-1 Resultados de la prueba de resistencia a la tensión de largueros de aluminio

Calidad del larguero	Tamaño de la muestra	Media muestral de la resistencia a la tensión (kg/mm^2)	Desviación estándar (kg/mm^2)
1	$n_1 = 10$	$\bar{x}_1 = 87.6$	$\sigma_1 = 1.0$
2	$n_2 = 12$	$\bar{x}_2 = 74.5$	$\sigma_2 = 1.5$

Elección del tamaño de la muestra

Si las desviaciones estándar σ_1 y σ_2 se conocen (al menos aproximadamente) y los dos tamaños de las muestras n_1 y n_2 son iguales ($n_1 = n_2 = n$, por ejemplo), entonces puede determinarse el tamaño de la muestra requerido para que el error al estimar $\mu_1 - \mu_2$ mediante $\bar{x}_1 - \bar{x}_2$, sea menor que E con una confianza del $100(1 - \alpha)$ por ciento. El tamaño de la muestra requerido en cada población es

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 (\sigma_1^2 + \sigma_2^2) \tag{9-10}$$

Recuérdese redondear hacia arriba n si no es un entero. Con esto se asegurará que el nivel de confianza no descenderá abajo del $100(1 - \alpha)$ por ciento.

Intervalos de confianza de una cola

También pueden obtenerse intervalos de confianza de una cola para $\mu_1 - \mu_2$. Un intervalo de confianza superior del $100(1 - \alpha)$ por ciento para $\mu_1 - \mu_2$ es

$$\mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{9-11}$$

y un intervalo de confianza inferior del $100(1 - \alpha)$ por ciento es

$$\bar{x}_1 - \bar{x}_2 - z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \tag{9-12}$$

EJERCICIOS DE LA SECCIÓN 9-2

9-1. Se usan dos máquinas para llenar botellas de plástico con un volumen neto de 16.0 onzas. Puede suponerse que el volumen de llenado es normal, con una desviación estándar de $\sigma_1 = 0.020$ y $\sigma_2 = 0.025$ onzas. Uno de los miembros del personal de ingeniería de calidad sospecha que ambas máquinas hacen el llenado con el mismo volumen neto medio, sea este volumen 16.0 onzas o no. Se toma una muestra aleatoria de 10 botellas de la producción de cada máquina.

Máquina 1		Máquina 2	
16.03	16.01	16.02	16.03
16.04	15.96	15.97	16.04
16.05	15.98	15.96	16.02
16.05	16.02	16.01	16.01
16.02	15.99	15.99	16.00

a) ¿Piensa el lector que el ingeniero está en lo correcto? Use $\alpha = 0.05$.
b) ¿Cuál es el valor P para esta prueba?
c) ¿Cuál es la potencia de la prueba del inciso a) para una diferencia real en las medias de 0.08?
d) Encuentre un intervalo de confianza de 95% para la diferencia en las medias. Proponga una interpretación práctica de este intervalo.
e) Suponiendo tamaños de las muestras iguales, ¿qué tamaño de la muestra deberá usarse para asegurarse de que $\beta = 0.05$ si la verdadera diferencia en las medias es0.08? Suponga que $\alpha = 0.05$.

9-2. Un fabricante de componentes electrónicos puede usar indistintamente dos tipos de plástico. La resistencia a la ruptura de este plástico es importante. Se sabe que $\sigma_1 = \sigma_2 = 1.0$ psi. A partir de muestras aleatorias de tamaño $n_1 = 10$ y $n_2 = 12$, se obtiene $\bar{x}_1 = 162.5$ y $\bar{x}_2 = 155.0$. La compañía no adoptará el plástico 1 a menos que su resistencia media a la ruptura exceda la del plástico 2 en al menos 10 psi. Con base en la información de las muestras, ¿deberá utilizar el plástico 1? Use $\alpha = 0.05$ para llegar a una decisión.

9-3. Considérese de nuevo la situación del ejercicio 9-2. Suponga que la verdadera diferencia en las medias es en realidad 12 psi. Encuentre la potencia de la prueba suponiendo que $\alpha = 0.05$. Si en realidad es importante detectar esta diferencia, en opinión del lector, ¿son adecuados los tamaños de las muestras usados en el ejercicio 9-2?

9-4. Se estudia la velocidad de combustión de dos cargas propulsoras sólidas diferentes usadas en el sistema de expulsión de la tripulación de un avión. Se sabe que ambas cargas propulsoras sólidas tienen aproximadamente la misma desviación estándar para la velocidad de combustión; esto es, $\sigma_1 = \sigma_2 = 3$ cm/s. Se prueban dos muestras aleatorias de $n_1 = 20$ y $n_2 = 20$ ejemplares; las medias muestrales de la velocidad de combustión son $\bar{x}_1 = 18$ cm/s y $\bar{x}_2 = 24$ cm/s.

a) Pruebe la hipótesis de que ambas cargas propulsoras tienen la misma velocidad de combustión media. Use $\alpha = 0.05$.
b) ¿Cuál es el valor P de la prueba del inciso a)?
c) ¿Cuál es el error β de la prueba del inciso a) si la verdadera diferencia de la velocidad de combustión media es 2.5 cm/s?
d) Construya un intervalo de confianza de 95% para la diferencia de las medias $\mu_1 - \mu_2$. ¿Cuál es el significado práctico de este intervalo?

9-5. Se usan dos máquinas para llenar botellas con detergente para lavar trastes. Se sabe que las desviaciones estándar del volumen de llenado son $\sigma_1 = 0.10$ onzas líquidas y $\sigma_2 = 0.15$ onzas líquidas, respectivamente, para las dos máquinas. Se seleccionan dos muestras aleatorias de tamaños $n_1 = 12$ botellas de la máquina 1 y $n_2 = 10$ botellas de la máquina 2, y los volúmenes medios de llenado muestrales son $\bar{x}_1 = 30.87$ onzas líquidas y $\bar{x}_2 = 30.68$ onzas líquidas. Suponga una distribución normal.

a) Construya un intervalo de confianza de dos colas de 90% para la diferencia en las medias del volumen de llenado. Interprete este intervalo.
b) Construya un intervalo de confianza de dos colas de 95% para la diferencia en las medias del volumen de llenado. Compare la anchura de este intervalo con la del intervalo del inciso a) y comente.
c) Construya un intervalo de confianza superior de 95% para la diferencia en las medias del volumen de llenado. Interprete este intervalo.

9-6. Considérese de nuevo la situación descrita en el ejercicio 9-5.

a) Pruebe la hipótesis de que ambas máquinas hacen el llenado con el mismo volumen medio. Use a = 0.05.
b) ¿Cuál es el valor P de la prueba del inciso a)?
c) Si el error β de la prueba cuando la verdadera diferencia del volumen de llenado es 0.2 onzas líquidas no deberá exceder 0.1, ¿que tamaños de las muestras deben utilizarse? Use a = 0.05.

9-7. Dos diferentes formulaciones de un combustible oxigenado para motores se prueban para estudiar sus índice de octano en carretera. La varianza del índice de octano en carretera para la formulación 1 es $\sigma_1^2 = 1.5$, y para la formulación 2 es $\sigma_2^2 = 1.2$. Se prueban dos muestras aleatorias de tamaño $n_1 = 15$ y $n_2 = 20$, y las medias del índice de octano en carretera observadas son $\bar{x}_1 = 89.6$ y $\bar{x}_2 = 92.5$. Suponga una distribución normal.
a) Construya un intervalo de confianza de dos colas de 95% para la diferencia de las medias del índice de octano en carretera.
b) Si la formulación 2 produce un índice de octano en carretera más alto que la formulación 1, al fabricante le gustaría detectarlo. Formule y pruebe una hipótesis apropiada usando $\alpha = 0.05$.
c) ¿Cuál es el valor P para la prueba realizada en el inciso b)?

9-8. Considérese la situación descrita en el ejemplo 9-4. ¿Qué tamaño de la muestra se necesitaría en cada población si se quisiera que el error al estimar la diferencia en las medias de la velocidad de combustión fuera menor que 4 cm/s con una confianza de 99%?

9-9. Considérese la situación de la prueba del índice de octano en carretera descrita en el ejercicio 9-7. ¿Qué tamaño de la muestra se necesitaría en cada población si se quiere tener una confianza de 95% de que el error al estimar la diferencia en las medias del índice de octano en carretera sea menor que 1?

9-10. Un polímero se fabrica en un proceso químico por lotes. Se hacen mediciones regulares de la viscosidad de cada lote, y una larga experiencia con el proceso indica que la variabilidad del proceso es bastante estable con $\sigma = 20$. Se presenta a continuación la medición de la viscosidad de 15 lotes: 724, 718, 776, 760, 745, 759, 795, 756, 742, 740, 761, 749, 739, 747, 742. Se hace una modificación en el proceso, la cual incluye el cambio del tipo de catalizador usado en el proceso. Después de la modificación del proceso, se hace la medición de la viscosidad en ocho lotes: 735, 775, 729, 755, 783, 760, 738, 780. Suponga que la variabilidad del proceso no es afectada por el cambio de catalizador. Encuentre un intervalo de confianza de 90% para la diferencia en la viscosidad media de los lotes que resulta de la modificación del proceso.

9-11. Se piensa que la concentración del ingrediente activo de un detergente líquido para lavar ropa es afectada por el tipo de catalizador que se utiliza en el proceso. Se sabe que la desviación estándar de la concentración activa es 3 g/l, independientemente del tipo de catalizador. Se hacen 10 observaciones de la concentración con cada catalizador, y los datos se presentan a continuación:

Catalizador 1: 57.9, 66.2, 65.4, 65.4, 65.2, 62.6, 67.6, 63.7, 67.2, 71.0

Catalizador 2: 66.4, 71.7, 70.3, 69.3, 64.8 69.6, 68.6, 69.4, 65.3, 68.8

a) Encuentre un intervalo de confianza de 95% para la diferencia en las medias de la concentración activa de los dos catalizadores.
b) ¿Hay evidencia que indique que las medias de la concentración activa dependen de la elección del catalizador? Apóyese en los resultados del inciso a) para responder.

9-12. Considérense los datos de la viscosidad del polímero de los lotes del ejercicio 9-10. Si la diferencia de la viscosidad media de los lotes es 10 o menos, al fabricante le gustaría detectarlo con una probabilidad alta.
a) Formule y pruebe una hipótesis apropiada usando $\alpha = 0.10$. ¿Cuáles son las conclusiones?

b) Calcule el valor P para esta prueba.

c) Compare los resultados de los inciso a) y b) con la longitud del intervalo de confianza de 90% obtenido en el ejercicio 9-10 y comente los resultados.

9-13. En el problema del detergente para lavar ropa del ejercicio 9-11, pruebe la hipótesis de que las medias de la concentración activa son iguales para ambos tipos de catalizador. Use $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba? Compare la respuesta con la que se encontró en el inciso b) del ejercicio 9-11, y comente por qué son iguales o diferentes.

9-14. Considérese de nuevo el problema del detergente para lavar ropa del ejercicio 9-11. Suponga que la verdadera diferencia de las medias de la concentración activa es 5 g/l. ¿Cuál es la potencia de la prueba para detectar esta diferencia si $\alpha = 0.05$? Si esta diferencia es realmente importante, ¿considera usted que los tamaños de las muestras usados en los experimentos son adecuados?

9-3 INFERENCIA PARA LA DIFERENCIA EN MEDIAS DE DOS DISTRIBUCIONES NORMALES, VARIANZAS DESCONOCIDAS

Se amplían ahora los resultados de la sección anterior a la diferencia en las medias de las dos distribuciones de la figura 9-1 cuando las varianzas de ambas distribuciones σ_1^2 y σ_2^2 son desconocidas. Si los tamaños de las muestras n_1 y n_2 exceden 30, entonces sería posible aplicar los procedimientos de la distribución normal de la sección 9-2. Sin embargo, cuando se trabaja con muestras pequeñas, se supondrá que las poblaciones tienen una distribución normal, y las pruebas de hipótesis y los intervalos de confianza se basarán en la distribución t. Se trata de un paralelo exactamente con el caso de la inferencia sobre la media de una sola muestra con varianza desconocida.

9-3.1 Pruebas de hipótesis para la diferencia en medias, varianzas desconocidas

Se consideran ahora la prueba de hipótesis sobre la diferencia en las medias $\mu_1 - \mu_2$ de dos distribuciones normales donde las varianzas σ_1^2 y σ_2^2 son desconocidas. Se usará un estadístico t para probar estas hipótesis. Como ya se señaló arriba y en la sección 8-3, se requiere el supuesto de normalidad para desarrollar el procedimiento de prueba, pero desviaciones moderadas de la normalidad no tienen efectos adversos sobre el procedimiento. Deben tratarse dos situaciones diferentes. En el primer caso, se supone que las varianzas de las dos distribuciones normales son desconocidas pero iguales; es decir, $\sigma_1^2 = \sigma_2^2 = \sigma^2$. En el segundo, se supone que σ_1^2 y σ_2^2 son desconocidas y no necesariamente iguales.

Caso 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Suponga que se tienen dos poblaciones normales independientes con medias μ_1 y μ_2 desconocidas, y varianzas desconocidas pero iguales $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Quiere probarse

$$\begin{aligned} H_0&: \mu_1 - \mu_2 = \Delta_0 \\ H_1&: \mu_1 - \mu_2 \neq \Delta_0 \end{aligned} \tag{9-13}$$

Sea $X_{11}, X_{12}, \ldots, X_{2n_1}$ una muestra aleatoria de n_1 observaciones de la primera población y $X_{21}, X_{22}, \ldots, X_{2n_2}$ una muestra aleatoria de n_2 observaciones de la segunda población. Sean

$\bar{X}_1, \bar{X}_2, S_1^2, S_2^2$ las medias muestrales y las varianzas muestrales, respectivamente. Entonces, el valor esperado de la diferencia de las medias muestrales $\bar{X}_1 - \bar{X}_2$ es $E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$, por lo que $\bar{X}_1 - \bar{X}_2$ es un estimador insesgado de la diferencia de las medias. La varianza de $\bar{X}_1 - \bar{X}_2$ es

$$V(\bar{X}_1 - \bar{X}_2) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Parece razonable combinar las dos varianzas muestrales S_1^2 y S_2^2 para formar un estimador de σ^2. El **estimador combinado** de σ^2 se define como sigue:

El **estimador combinado** de σ^2, denotado por S_p^2, está definido por

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \tag{9-14}$$

Es sencillo ver que el estimador combinado S_p^2 puede escribirse como

$$\begin{aligned} S_p^2 &= \frac{n_1 - 1}{n_1 + n_2 - 2} S_1^2 + \frac{n_2 - 1}{n_1 + n_2 - 2} S_2^2 \\ &= wS_1^2 + (1 - w)S_2^2 \end{aligned}$$

donde $0 < w \leq 1$. Por tanto, S_p^2 es un **promedio ponderado** de las dos varianzas muestrales S_1^2 y S_2^2, donde las ponderaciones w y $1 - w$ dependen de los dos tamaños de las muestras n_1 y n_2. Obviamente, si $n_1 = n_2 = n$, entonces $w = 0.5$ y S_p^2 es simplemente el promedio aritmético de S_1^2 y S_2^2. Si $n_1 = 10$ y $n_2 = 20$ (por ejemplo), entonces $w = 0.32$ y $1 - w = 0.68$. La primera muestra contribuye con $n_1 - 1$ grados de libertad para S_p^2 y la segunda muestra contribuye con $n_2 - 1$ grados de libertad. Por lo tanto, S_p^2 tiene $n_1 + n_2 - 2$ grados de libertad.

Se sabe ahora que

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

tiene una distribución $N(0, 1)$. Al reemplazar σ con S_p, se llega al siguiente resultado:

Dados los supuestos de esta sección, la cantidad

$$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{9-15}$$

tiene una distribución t con $n_1 + n_2 - 2$ grados de libertad.

Ahora es directo el uso de esta información para probar las hipótesis de la ecuación 9-13: simplemente se reemplaza $\mu_1 - \mu_2$ con Δ_0, y el **estadístico de la prueba** resultante tiene una distribución t con $n_1 + n_2 - 2$ grados de libertad bajo H_0: $\mu_1 - \mu_2 = \Delta_0$. La localización de la región crítica para hipótesis alternativas tanto de dos colas como de una cola es semejante al caso de una sola muestra. Puesto que se usa una estimación combinada de la varianza, el proceso suele conocerse como la prueba t combinada.

La prueba t combinada para dos muestras[1]

Hipótesis nula: H_0: $\mu_1 - \mu_2 = \Delta_0$

Estadístico de la prueba:

$$T_0 = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{9-16}$$

Hipótesis alternativa	Criterio de rechazo
H_1: $\mu_1 - \mu_2 \neq \Delta_0$	$t_0 > t_{\alpha/2, n_1+n_2-2}$ o $t_0 < -t_{\alpha/2, n_1+n_2-2}$
H_1: $\mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha, n_1+n_2-2}$
H_1: $\mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha/2, n_1+n_2-2}$

[1] Aun cuando se ha presentado el desarrollo de este procedimiento para el caso en que los tamaños de las muestras podrían ser diferentes, tiene una ventaja usar tamaños de las muestras iguales $n_1 = n_2 = n$. Cuando los tamaños de las muestras son iguales en ambas poblaciones, la prueba t es muy robusta en cuanto al supuesto de varianzas iguales.

Tabla 9-2 Datos del rendimiento de los catalizadores del ejemplo 9-5

Número de observación	Catalizador 1	Catalizador 2
1	91.50	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75
	$\bar{x}_1 = 92.255$	$\bar{x}_2 = 92.733$
	$s_1 = 2.39$	$s_2 = 2.98$

EJEMPLO 9-5

Se analizan dos catalizadores para determinar la forma en que afectan el rendimiento medio de un proceso químico. Específicamente, el catalizador 1 se encuentra en uso actualmente, pero el catalizador 2 es aceptable. Puesto que el catalizador 2 es más barato, debería adoptarse, siempre que no modifique el rendimiento del proceso. Se realiza una prueba en la planta piloto y en la tabla 9-2 se dan los resultados. ¿Hay alguna diferencia entre los rendimientos medios? Use $\alpha = 0.05$, y suponga varianzas iguales.

La solución usando el procedimiento de prueba de hipótesis de ocho pasos es la siguiente:

1. Los parámetros de interés son μ_1 y μ_2, el rendimiento medio del proceso usando los catalizadores 1 y 2, respectivamente, y quiere saberse si $\mu_1 - \mu_2 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, o H_0: $\mu_1 = \mu_2$
3. H_1: $\mu_1 \neq \mu_2$
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

6. Se rechaza H_0 si $t_0 > t_{0.025,\,14} = 2.145$ o si $t_0 < -t_{0.025,\,14} = -2.145$.
7. Cálculos: Por la tabla 9-2 se tiene $\bar{x}_1 = 92.255$, $s_1 = 2.39$, $n_1 = 8$, $\bar{x}_2 = 92.733$, $s_2 = 2.98$ y $n_2 = 8$. Por lo tanto,

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(7)(2.39)^2 + 7(2.98)^2}{8 + 8 - 2} = 7.30$$

$$s_p = \sqrt{7.30} = 2.70$$

y

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{2.70\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{92.255 - 92.733}{2.70\sqrt{\frac{1}{8} + \frac{1}{8}}} = -0.35$$

8. Conclusiones: Puesto que $-2.145 < t_0 = -0.35 < 2.145$, no puede rechazarse la hipótesis nula. Es decir, con el nivel de significación 0.05 no se cuenta con evidencia sólida para concluir que el catalizador 2 produce un rendimiento medio que difiere del rendimiento medio cuando se usa el catalizador 1.

También pudo usarse un valor P para tomar una decisión en este ejemplo. En la tabla IV del apéndice se encuentra que $t_{0.40,\,14} = 0.258$ y $t_{0.25,\,14} = 0.692$. Por lo tanto, puesto que $0.258 < 0.35 < 0.692$, se concluye que los límites inferior y superior del valor P son $0.50 < P < 0.80$. De hecho, el valor real es $P = 0.7315$. (Este valor se obtuvo con una calculadora manual.) Por lo tanto, puesto que el valor P excede $\alpha = 0.05$, la hipótesis nula no puede rechazarse.

Caso 2: $\sigma_1^2 \neq \sigma_2^2$

En algunas situaciones, no es razonable suponer que las varianzas desconocidas σ_1^2 y σ_2^2 son iguales. No hay un estadístico t exacto para probar H_0: $\mu_1 - \mu_2 = \Delta_0$ en este caso. Sin embargo, si H_0: $\mu_1 - \mu_2 = \Delta_0$ es verdadera, entonces el estadístico

$$T_0^* = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \tag{9-17}$$

tiene una distribución aproximadamente igual a la de t con los grados de libertad dados por

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(S_1^1/n_1\right)^2}{n_1+1} + \frac{\left(S_2^2/n_2\right)^2}{n_2+1}} - 2 \tag{9-18}$$

si la hipótesis nula H_0: $\mu_1 - \mu_2 = \Delta_0$ es verdadera. Por lo tanto, si $\sigma_1^2 \neq \sigma_2^2$, las hipótesis sobre las diferencias en las medias de dos distribuciones normales se prueban como en el caso de varianzas iguales, excepto porque se usa T_0^* como el estadístico de la prueba y $n_1 + n_2 - 2$ se reemplaza con v para determinar los grados de libertad de la prueba.

EJEMPLO 9-6

Un fabricante de unidades reproductoras de video está probando el diseño de dos microcircuitos para determinar si producen un flujo de corriente medio equivalente. El departamento de ingeniería de desarrollo ha obtenido los siguientes datos:

Diseño 1:	$n_1 = 5$	$\bar{x}_1 = 24.2$	$s_1^2 = 10$
Diseño 2:	$n_2 = 10$	$\bar{x}_2 = 23.9$	$s_2^2 = 20$

Usando $\alpha = 0.10$, quiere determinarse si hay alguna diferencia en el flujo de corriente medio entre los dos diseños, donde se supone que ambas poblaciones son normales, pero no se desea suponer que las varianzas desconocidas σ_1^2 y σ_2^2 son iguales.

Al aplicar el procedimiento de ocho pasos se obtiene:

1. Los parámetros de interés son los flujos de corriente medios del diseño de los dos circuitos, por ejemplo, μ_1 y μ_2, y el interés se encuentra en determinar si $\mu_1 - \mu_2 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, o H_0: $\mu_1 = \mu_2$
3. H_1: $\mu_1 \neq \mu_2$
4. $\alpha = 0.10$
5. El estadístico de la prueba es

$$t_0^* = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

6. Por la ecuación 9-18 se encuentra que los grados de libertad sobre t_0^* son

$$v = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_1^2/n_1)^2}{n_1+1} + \dfrac{(s_2^2/n_2)^2}{n_2+1}} - 2 = \frac{\left(\dfrac{10}{15} + \dfrac{20}{10}\right)^2}{\dfrac{(10/15)^2}{16} + \dfrac{(20/10)^2}{11}} - 2 = 16.17 \simeq 16$$

Por lo tanto, puesto que $\alpha = 0.10$, se rechazaría H_0: $\mu_1 = \mu_2$ si $t_0^* > t_{0.05,\,16} = 1.746$ o si $t_0^* < -t_{0.05,\,16} = -1.746$.

7. Cálculos: Al usar los datos muestrales se obtiene

$$t_0^* = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{24.2 - 23.9}{\sqrt{\frac{10}{15} + \frac{20}{10}}} = 0.18$$

8. Conclusiones: Puesto que $-1.746 < t_0^* = 0.18 < 1.746$, no puede rechazarse H_0: $\mu_1 = \mu_2$ con el nivel de significación $\alpha = 0.10$. Es decir, no hay evidencia sólida que indique que el flujo de corriente medio es diferente para los dos diseños. El valor P para $t_0^* = 0.18$ es aproximadamente 0.859.

9-3.2 Elección del tamaño de la muestra

Se usan las curvas de operación característica de las gráficas VI*e*, VI*f*, VI*g* y VI*h* del apéndice para evaluar el error tipo II en el caso en que $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Desafortunadamente, cuando $\sigma_1^2 \neq \sigma_2^2$, no se conoce la distribución de T_0^* si la hipótesis nula es falsa, y no se cuenta con curvas de operación característica para este caso.

Para la hipótesis alternativa de dos colas H_1: $\mu_1 - \mu_2 = \Delta \neq \Delta_0$, cuando $\sigma_1^2 = \sigma_2^2 = \sigma^2$ y $n_1 = n_2 = n$, se usan las gráficas VI*e* y VI*f* con

$$d = \frac{|\Delta - \Delta_0|}{2\sigma} \tag{9-19}$$

donde Δ es la verdadera diferencia en las medias que es de interés. Para usar estas curvas, deben consultarse con el tamaño de la muestra $n^* = 2n - 1$. Para la hipótesis alternativa de una cola se usan las gráficas VI*g* y VI*h*, y d y Δ se definen como en la ecuación 9-19. Se observa que el parámetro d es una función de σ, la cual es desconocida. Como en una prueba t para una sola muestra, quizás haya que apoyarse en una estimación previa de σ o usar una estimación subjetiva. De manera alternativa, podrían definirse las diferencias en la media que quieren detectarse con relación a σ.

EJEMPLO 9-7

Considérese el experimento del catalizador del ejemplo 9-5. Suponga que, si el catalizador 2 produce un rendimiento medio que difiere del rendimiento medio del catalizador 1 por 4.0%, querría rechazarse la hipótesis nula con una probabilidad de al menos 0.85. ¿Qué tamaño de la muestra se necesita?

Usando $s_p = 2.70$ como estimación aproximada de la desviación estándar común σ, se tiene $d = |\Delta|/2\sigma = |4.0|/[(2)(2.70)] = 0.74$. Por la gráfica VI*e* del apéndice, con $d = 0.74$ y $\beta = 0.15$, se encuentra $n^* = 20$, aproximadamente. Por lo tanto, puesto que $n^* = 2n - 1$,

$$n = \frac{n^* + 1}{2} = \frac{20 + 1}{2} = 10.5 \simeq 11 \text{ (por decir así)}$$

y se usarían los tamaños de las muestras $n_1 = n_2 = n = 11$.

9-3.3 Intervalo de confianza para la diferencia en medias

Caso 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Para desarrollar el intervalo de confianza para la diferencia en medias $\mu_1 - \mu_2$ cuando las dos varianzas son iguales, se observa que la distribución del estadístico

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

es la distribución t con $n_1 + n_2 - 2$ grados de libertad. Por lo tanto,

$$P\left(-t_{\alpha/2,\, n_1+n_2-2} \le T \le t_{\alpha/2,\, n_1+n_2-2}\right) = 1 - \alpha$$

o bien

$$P\left(-t_{\alpha/2,n_1+n_2-2} \le \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \le t_{\alpha/2,\, n_1-n_2-2}\right) = 1 - \alpha$$

La manipulación de las cantidades en el interior del enunciado de probabilidad lleva al intervalo de confianza del $100(1 - \alpha)$ por ciento para $\mu_1 - \mu_2$.

Definición: Intervalo de confianza para la diferencia en las medias de dos distribuciones normales, varianzas desconocidas pero iguales

Si $\bar{x}_1$, $\bar{x}_2$, s_1^2 y s_2^2 son las medias y las varianzas muestrales de dos muestras aleatorias de tamaños n_1 y n_2, respectivamente, de dos poblaciones normales independientes con varianzas desconocidas pero iguales, entonces un intervalo de confianza del $100(1 - \alpha)$ por ciento para la diferencia en las medias $\mu_1 - \mu_2$ es

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha/2,n_1+n_2-2}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{\alpha/2,\, n_1+n_2-2}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad (9\text{-}20)$$

donde $s_p = \sqrt{[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2]/(n_1 + n_2 - 2)}$ es la estimación combinada de la desviación estándar común poblacional, y $t_{\alpha/2,n_1+n_2-2}$ es el punto porcentual $\alpha/2$ superior de la distribución t con $n_1 + n_2 - 2$ grados de libertad.

EJEMPLO 9-8

Un artículo de la revista *Waste Hazardous and Hazardous Materials* (Vol. 6) reportó los resultados de un análisis del peso del calcio en el cemento estándar y el cemento adicionado con plomo. Niveles reducidos de calcio indicarían que se bloquea el mecanismo de hidratación del cemento, lo cual permitiría que el agua atacara varios lugares de la estructura del cemento. Diez muestras de cemento estándar tuvieron un peso promedio porcentual de calcio de $\bar{x}_1 = 90.0$, con una desviación estándar muestral de $s_1 = 5.0$, mientras que 15 muestras de cemento adicionado con plomo tuvieron un peso promedio porcentual de calcio de $\bar{x}_2 = 87.0$, con una desviación estándar muestral de $s_2 = 4.0$.

Se supondrá que el peso porcentual de calcio tiene una distribución normal y se encontrará un intervalo de confianza de 95% para la diferencia en las medias, $\mu_1 - \mu_2$, para los dos tipos de cemento. Además, se supondrá que ambas poblaciones normales tienen la misma desviación estándar.

La estimación combinada de la desviación estándar común se encuentra usando la ecuación 9-14 como sigue:

$$\begin{aligned} s_p^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \\ &= \frac{9(5.0)^2 + 14(4.0)^2}{10 + 15 - 2} \\ &= 19.52 \end{aligned}$$

Por lo tanto, la estimación combinada de la desviación estándar es $s_p = \sqrt{19.52} = 4.4$. El intervalo de confianza de 95% se encuentra aplicando la ecuación 9-20:

$$\bar{x}_1 - \bar{x}_2 - t_{0.025,23} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{0.025,23} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

o al sustituir los valores muestrales y usar $t_{0.025,23} = 2.069$,

$$\begin{aligned} 90.0 - 87.0 - 2.069(4.4)\sqrt{\frac{1}{10} + \frac{1}{15}} &\le \mu_1 - \mu_2 \\ &\le 90.0 - 87.0 + 2.069(4.4)\sqrt{\frac{1}{10} + \frac{1}{15}} \end{aligned}$$

expresión que se reduce a

$$-0.72 \le \mu_1 - \mu_2 \le 6.72$$

Obsérvese que el intervalo de confianza de 95% incluye a cero; por lo tanto, con este nivel de confianza no puede concluirse que hay una diferencia en las medias. Dicho de otra manera, no hay evidencia de que el cemento adicionado con plomo afectara el peso promedio porcentual de calcio; por lo tanto, no puede afirmarse que la presencia de plomo afecta este aspecto del mecanismo de hidratación con el nivel de confianza de 95%.

Caso 2: $\sigma_1^2 \neq \sigma_2^2$

En muchas situaciones no es razonable suponer que $\sigma_1^2 = \sigma_2^2$. Cuando este supuesto no está garantizado, aún puede encontrarse un intervalo de confianza del 100(1 − α) por ciento para $\mu_1 - \mu_2$ utilizando el hecho de que

$$T^* = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

tiene una distribución aproximadamente igual a la de t con v grados de libertad dada por la ecuación 9-18. Por lo tanto,

$$P\left(-t_{\alpha/2,v} \leq T^* \leq t_{\alpha/2,v}\right) \cong 1 - \alpha$$

y si se sustituye T^* en esta expresión y se aísla $\mu_1 - \mu_2$ entre las desigualdades, puede obtenerse el intervalo de confianza para $\mu_1 - \mu_2$.

Definición: Intervalo de confianza para la diferencia en las medias de dos distribuciones normales, varianzas desconocidas y diferentes

Si $\bar{x}_1, \bar{x}_2, s_1^2$ y s_2^2 son las medias y las varianzas de dos muestras aleatorias de tamaños n_1 y n_2, respectivamente, de dos poblaciones normales independientes con varianzas desconocidas y diferentes, entonces un intervalo de confianza del 100(1 − α) por ciento aproximado para la diferencia de las medias $\mu_1 - \mu_2$ es

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad (9\text{-}21)$$

donde v está dada por la ecuación 9-18 y $t_{\alpha/2,v}$ es el punto porcentual α/2 superior de la distribución t con v grados de libertad.

9-3.4 Solución de computadora

La prueba t para dos muestras puede efectuarse utilizando la mayoría de los paquetes de software estadístico. En la tabla 9-3 se presenta la salida de la rutina de la prueba t para dos muestras del programa Minitab para los datos del rendimiento del catalizador del ejemplo 9-5. La salida incluye resumen de estadísticos para cada muestra, intervalos de confianza para la diferencia de las medias y los resultados de la prueba de hipótesis. Este análisis se realizó suponiendo varianzas

Tabla 9-3 Salida del programa Minitab de la prueba *t* para dos muestras del ejemplo 9-5

Prueba *t* para dos muestras e intervalo de confianza

Catalizador 1	Catalizador 2
91.50	89.19
94.18	90.95
92.18	90.46
95.39	93.21
91.79	97.19
89.07	97.04
94.72	91.07
89.21	92.75

Dos muestras *t* para el catalizador 1 contra el catalizador 2

	N	Media	Desv. Est.	EE Media
Catalizador 1	8	92.26	2.39	0.84
Catalizador 2	8	92.73	2.98	1.1

95% CI para el catalizador 1 mu – catalizador 2 mu: (–3.39, 2.4)
Prueba *t* mu catalizador 1 = mu catalizador 2 (en contra no es =) : T = –0.35 P = 0.73 DF = 13

iguales. Minitab tiene una opción para realizar el análisis suponiendo varianzas diferentes. El usuario puede especificar los niveles de confianza y el valor α. El procedimiento de prueba de hipótesis indica que no puede rechazarse la hipótesis de que los rendimientos medios son iguales, lo cual concuerda con las conclusiones a las que se llegó originalmente en el ejemplo 9-5.

En la figura 9-2 se muestran las gráficas de caja comparativas para los datos del rendimiento de los dos tipos de catalizadores del ejemplo 9-5. Estas gráficas de caja comparativas indican que no hay ninguna diferencia obvia en las medianas de las dos muestras, aun cuando la segun-

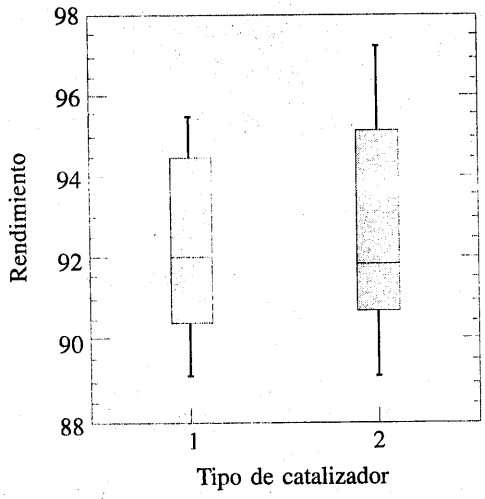


Figura 9-2 Gráficas de caja comparativas para los datos del rendimiento de los catalizadores del ejemplo 9-5.

da muestra tiene una dispersión o varianza ligeramente mayor. No hay reglas exactas para comparar dos muestras con gráficas de caja; su valor principal radica en la impresión visual que proporcionan como herramienta para explicar los resultados de una prueba de hipótesis, así como en la verificación de los supuestos.

En la figura 9-3 se presenta una gráfica de probabilidad normal de las dos muestras del ejemplo 9-5. Obsérvese que las muestras se grafican aproximadamente sobre líneas rectas, y que las rectas de cada muestra tienen pendientes similares. Por lo tanto, se concluye que los supuestos de normalidad y varianzas iguales son razonables.

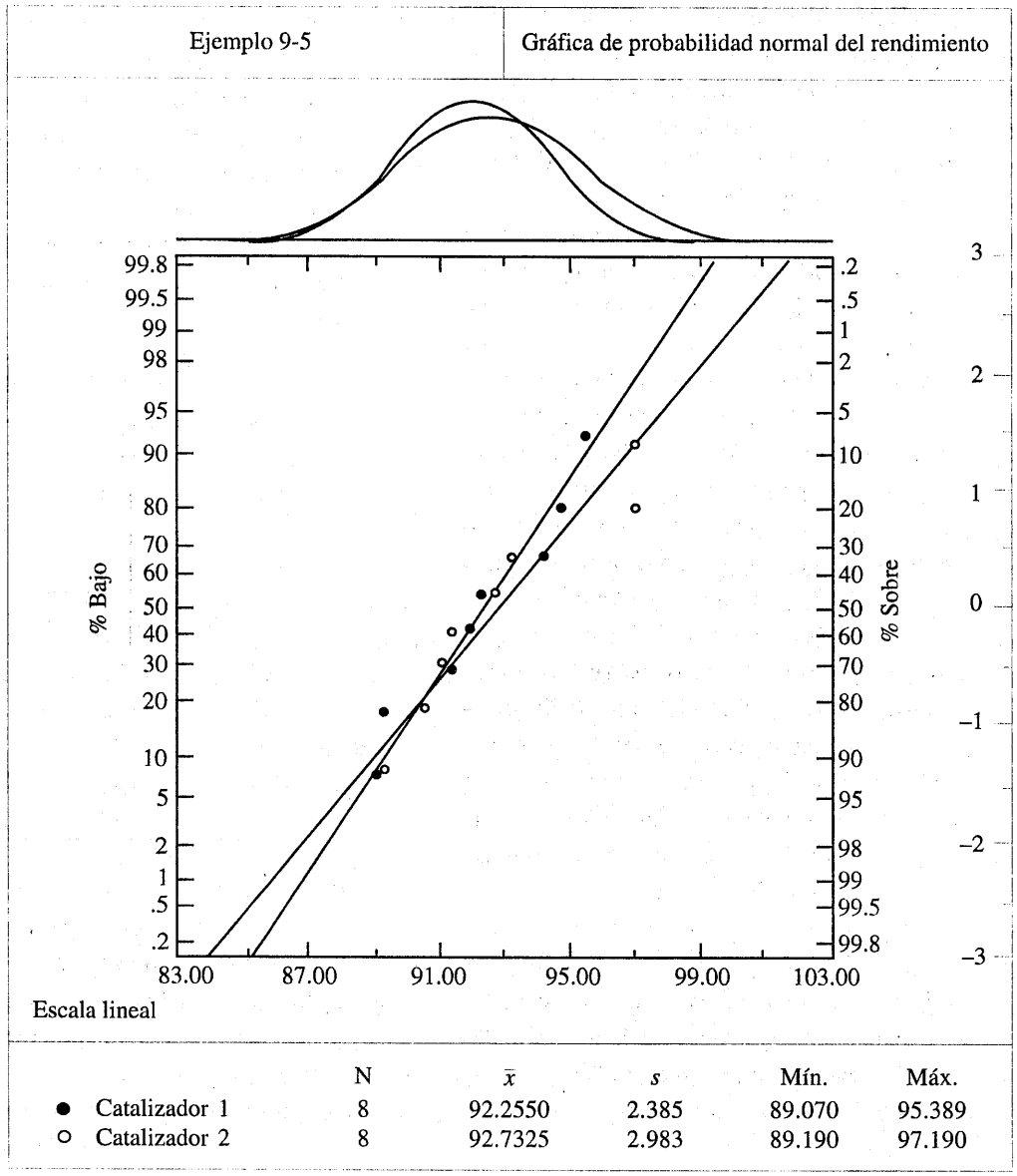


Figura 9-3 Gráfica de probabilidad normal para los datos del rendimiento del ejemplo 9-5.

EJERCICIOS DE LA SECCIÓN 9-3

9-15. Se investiga el diámetro de las varillas de acero fabricadas en dos máquinas de extrusión diferentes. Se seleccionan dos muestras aleatorias de tamaños $n_1 = 15$ y $n_2 = 17$, y las medias muestrales y las varianzas muestrales son $\bar{x}_1 = 8.73$, $s_1^2 = 0.35$, $\bar{x}_2 = 8.68$ y $s_2^2 = 0.40$, respectivamente. Suponga que $\sigma_1^2 = \sigma_2^2$ y que los datos se extraen de una distribución normal.

a) ¿Hay evidencia en favor de la afirmación de que las dos máquinas producen varillas con medias del diámetro diferentes? Use $\alpha = 0.05$ para llegar a esta conclusión.

b) Encuentre el valor P para el estadístico t calculado en el inciso a).

c) Construya un intervalo de confianza de 95% para la diferencia de las medias del diámetro de las varillas. Interprete este intervalo.

9-16. En un artículo de *Fire Technology* se investigaron dos diferentes agentes de expansión de espuma que pueden usarse en las boquillas del equipo aspersor para combatir incendios. Una muestra aleatoria de cinco observaciones con una espuma acuosa que forma una película (AFFF) tuvo una media muestral de 4.7 y una desviación estándar de 0.6. Una muestra aleatoria de cinco observaciones con concentrados del tipo alcohol (ATC) tuvieron una media muestral de 6.9 y una desviación estándar de 0.8. Encuentre un intervalo de confianza de 95% para la diferencia de la expansión media de la espuma de estos dos agentes. ¿Puede llegarse a alguna conclusión sobre cuál de los agentes produce la expansión media de la espuma más alta? Suponga que ambas poblaciones están representadas adecuadamente por distribuciones normales con la misma desviación estándar.

9-17. Pueden usarse dos catalizadores en un proceso químico por lotes. Se prepararon 12 lotes usando el catalizador 1, obteniéndose un rendimiento promedio de 86 y una desviación estándar muestral de 3. Se prepararon 15 lotes usando el catalizador 2, obteniéndose un rendimiento promedio de 89 con una desviación estándar de 2. Suponga que las mediciones del rendimiento tienen una distribución aproximadamente normal con la misma desviación estándar.

a) ¿Hay evidencia en favor de la afirmación de que el catalizador 2 produce un rendimiento medio más alto que el catalizador 1? Use $\alpha = 0.01$.

b) Encuentre un intervalo de confianza de 95% para la diferencia en los rendimientos promedio.

9-18. Se investiga la temperatura de deflexión bajo carga de dos tipos diferentes de tubos de plástico. Se prueban dos muestras aleatorias de 15 ejemplares, y las temperaturas de deflexión observadas se reportan a continuación (en °F):

Tipo 1		
206	193	192
188	207	210
205	185	194
187	189	178
194	213	205

Tipo 2		
177	176	198
197	185	188
206	200	189
201	197	203
180	192	192

a) Construya gráficas de caja y gráficas de probabilidad normal para las dos muestras. ¿Estas gráficas apoyan los supuestos de normalidad y varianzas iguales? Escriba una interpretación práctica de estas gráficas.

b) ¿Los datos apoyan la afirmación de que la temperatura de deflexión bajo carga del tubo tipo 2 excede la del tipo 1? Use $\alpha = 0.05$ para llegar a una conclusión.

c) Calcule el valor P para la prueba del inciso a).

d) Suponga que si la temperatura de deflexión media del tubo tipo 2 excede la del tipo 1 hasta por 5°F, es importante detectar esta diferencia con una probabilidad de al menos 0.90. ¿La elección de $n_1 = n_2 = 15$ en el inciso a) de este problema es adecuada?

9-19. En la manufactura de semiconductores, es común el uso de un proceso de grabado por remojo químico para eliminar el silicio de la parte posterior de las obleas antes de la metalización. La rapidez de grabado es una característica importante en este proceso y se sabe que sigue una distribución normal. Se compararon dos soluciones de grabado diferentes, usando dos muestras aleatorias de 10 obleas para cada solución. La rapidez de grabado observadas fueron las siguientes (en milipulgadas/min):

Solución 1		Solución 2	
9.9	10.6	10.2	10.0
9.4	10.3	10.6	10.2
9.3	10.0	10.7	10.7
9.6	10.3	10.4	10.4
10.2	10.1	10.5	10.3

a) ¿Los datos apoyan la afirmación de que la rapidez de grabado medio es la misma para ambas soluciones? Para llegar a una conclusión, use $\alpha = 0.05$ y suponga que las varianzas de las dos poblaciones son iguales.

b) Calcule el valor P para la prueba del inciso a).

c) Encuentre un intervalo de confianza de 95% para la diferencia de las medias de la rapidez de grabado.

d) Construya gráficas de probabilidad normal para las dos muestras. ¿Estas gráficas brindan apoyo a los supuestos de normalidad y varianzas iguales? Escriba una interpretación práctica de estas gráficas.

9-20. Dos proveedores fabrican un engrane de plástico usado en una impresora láser. La resistencia a impactos de estos engranes, medida en libras-pie, es una característica importante. Una muestra aleatoria de 10 engranes del proveedor 1 da como resultado $\bar{x}_1 = 290$ y $s_1 = 12$, mientras que otra muestra aleatoria de 16 engranes del segundo proveedor da como resultado $\bar{x}_2 = 321$ y $s_2 = 22$.

a) ¿Hay evidencia en favor de la afirmación de que los engranes del proveedor 2 tienen una resistencia a impactos media más alta? Use $\alpha = 0.05$ y suponga que las dos poblaciones tienen una distribución normal pero las varianzas no son iguales.

b) ¿Cuál es el valor P para esta prueba?

c) ¿Los datos apoyan la afirmación de que la resistencia a impactos media de los engranes del proveedor 2 es al menos 25 libras-pie más alta que la del proveedor 1? Haga los mismos supuestos que en el inciso a).

9-21. Considérese de nuevo la situación del inciso a) del ejercicio 9-20. Construya una estimación del intervalo de confianza para la diferencia en las medias de la resistencia a impactos, y explique cómo podría usarse este intervalo para responder la pregunta planteada respecto de las diferencias entre un proveedor y el otro.

9-22. Una película fotoconductora se fabrica con un espesor nominal de 25 milipulgadas. El ingeniero del producto quiere incrementar la rapidez media de la película, y piensa que puede conseguirlo reduciendo el espesor de la película a 20 milipulgadas. Se fabrican ocho muestras con cada espesor de la película en un proceso de producción piloto, y se mide la rapidez de la película (en $\mu J/pulg^2$). Para la película de 25 milipulgadas, los datos muestrales dieron como resultado $\bar{x}_1 = 1.15$ y $s_1 = 0.11$, mientras que para la película de 20 milipulgadas los datos dieron como resultado $\bar{x}_2 = 1.06$ y $s_2 = 0.09$. Obsérvese que un incremento de la rapidez de la película reduciría el valor de la observación en $\mu J/pulg^2$.

a) ¿Los datos apoyan la afirmación de que reducir el espesor de la película incrementa la rapidez media de la película? Use $\alpha = 0.10$ y suponga que las dos varianzas poblacionales son iguales y que la población fundamental de la rapidez de la película tiene una distribución normal.

b) ¿Cuál es el valor P para esta prueba?

c) Encuentre un intervalo de confianza de 95% para la diferencia de las dos medias.

9-23. Los puntos de fusión de dos aleaciones usadas en la formulación de soldadura se investigaron fundiendo 21 muestras de cada material. La media y la desviación estándar muestrales de la aleación 1 fueron $\bar{x}_1 = 420°F$ y $s_1 = 4°F$, mientras para la aleación 2 fueron $\bar{x}_2 = 426°F$ y $s_2 = 3°F$. ¿Los datos muestrales apoyan la afirmación de que ambas aleaciones tienen el mismo punto de fusión? Use $\alpha = 0.05$ y suponga que ambas poblaciones tienen una distribución normal y la misma desviación estándar. Encuentre el valor P para la prueba.

9-24. Con referencia al experimento del punto de fusión del ejercicio 9-23, suponga que la verdadera diferencia en las medias del punto de fusión es 3°F. ¿Qué tamaño necesitaría tener una muestra para detectar esta diferencia usando un nivel de prueba $\alpha = 0.05$ con una probabilidad de al menos 0.9? Use $\sigma_1 = \sigma_2 = 4$ como estimación inicial de la desviación estándar común.

9-25. Dos compañías fabrican un material de hule que se usará en una aplicación automotriz. La pieza estará sometida a desgaste abrasivo en la aplicación de campo, por lo que se decide comparar el material producido por cada compañía en una prueba. Se prueban 25 muestras de material de cada compañía en una prueba de abrasión, y se observa la cantidad de desgaste después de 1 000 ciclos. Para la compañía 1, la media y la desviación estándar muestrales del desgaste son $\bar{x}_1 = 20$ mg/1 000 ciclos y $s_1 = 2$ mg/1 000 ciclos, mientras que para la compañía 2 se obtienen $\bar{x}_2 = 15$ mg/1 000 ciclos y $s_2 = 8$ mg/1 000 ciclos.

a) ¿Los datos apoyan la afirmación de que las dos compañías producen materiales con desgaste medio diferente? Use $\alpha = 0.05$, y suponga que cada población tiene una distribución normal pero que sus varianzas no son iguales.

b) ¿Cuál es el valor P para esta prueba?

c) ¿Los datos apoyan la afirmación de que el material de la compañía 1 tiene desgaste medio mayor que el material de la compañía 2? Use los mismos supuestos que en el inciso a).

9-26. Se piensa que en el espesor de una película de plástico (en milipulgadas) que se coloca sobre un material de sustrato influye la temperatura a la que se aplica el recubrimiento. Se lleva a cabo un experimento totalmente aleatorizado. Once sustratos se recubren a 125°F, dando como resultado un espesor medio muestral del recubrimiento de $\bar{x}_1 = 103.5$ y una desviación estándar muestral de $s_1 = 10.2$. Otros 13 sustratos se recubren a 150°F, para los que se observan $\bar{x}_2 = 99.7$ y $s_2 = 20.1$. Se creía al principio que con el incremento de la temperatura del proceso se reduciría el espesor medio del recubrimiento. ¿Los datos apoyan esta afirmación? Use $\alpha = 0.01$ y suponga que las desviaciones estándar de las dos poblaciones son diferentes. Calcule el valor P para esta prueba.

9-27. Considérese de nuevo el experimento del espesor del recubrimiento del ejercicio 9-26. ¿Cómo podría responderse la pregunta planteada respecto del efecto de la temperatura sobre el espesor del recubrimiento utilizando un intervalo de confianza? Explique su respuesta.

9-28. Considérese de nuevo la prueba del desgaste abrasivo del ejercicio 9-25. Construya un intervalo de confianza con el que puedan responderse las preguntas de los incisos a) y c) de ese ejercicio.

9-4 PRUEBA t PAREADA

Un caso especial de la prueba t para dos muestras de la sección 9-3 ocurre cuando las observaciones de las dos poblaciones de interés se colectan en **pares**. Cada par de observaciones, por ejemplo (X_{1j}, X_{2j}), se toma bajo condiciones homogéneas, pero estas condiciones pueden cam-

biar de un par a otro. Por ejemplo, suponga que el interés se encuentra en comparar dos tipos diferentes de puntas de una máquina para probar la dureza. Esta máquina presiona la punta en una ejemplar de prueba de metal con una fuerza conocida. Al medir la profundidad de la depresión producida por la punta puede determinarse la dureza del ejemplar de prueba. Si se seleccionaron varios ejemplares de prueba al azar, la mitad de los cuales se probaron con la punta 1 y la otra con la punta 2, y se aplicó la prueba t combinada o independiente de la sección 9-3, los resultados de la prueba podrían ser erróneos. Los ejemplares de prueba metálicos pudieron haberse cortado de barras de acero que se fabricaron a temperaturas diferentes, o podrían no ser homogéneos de alguna otra forma que afectara la dureza. Entonces la diferencia observada entre las lecturas de la dureza media para los dos tipos de punta también incluiría diferencias de dureza entre los ejemplares de pruebas.

Un procedimiento experimental más poderoso es colectar los datos en pares, es decir, hacer dos lecturas de dureza en cada ejemplar de prueba, una con cada punta. El procedimiento de prueba consistiría entonces en analizar las *diferencias* entre las lecturas de dureza de cada ejemplar de prueba. Si no hay diferencia entre las puntas, entonces la media de las diferencias deberá ser cero. A este procedimiento de prueba se le llama la **prueba *t* pareada**.

Sea (X_{11}, X_{21}), (X_{12}, X_{22}), ..., (X_{1n}, X_{2n}) un conjunto de n observaciones pareadas donde se supone que la media y la varianza de la población representada por X_1 son μ_1 y σ_1^2, y la media y la varianza de la población representada por X_2 son μ_2 y σ_2^2. Se definen las diferencias entre cada par de observaciones como $D_j = X_{1j} - X_{2j}$, $j = 1, 2, ..., n$. Se supone que las D_j tienen una distribución normal con media

$$\mu_D = E(X_1 - X_2) = E(X_1) - E(X_2) = \mu_1 - \mu_2$$

y varianza σ_D^2, por lo que probar hipótesis acerca de la diferencia entre μ_1 y μ_2 puede hacerse efectuando una prueba t de una muestra para μ_D. Específicamente, probar H_0: $\mu_1 - \mu_2 = \Delta_0$ contra H_1: $\mu_1 - \mu_2 \neq \Delta_0$ es equivalente a probar

$$\begin{aligned} H_0 &: \mu_D = \Delta_0 \\ H_1 &: \mu_D \neq \Delta_0 \end{aligned} \tag{9-22}$$

El estadístico de la prueba se da a continuación.

La prueba *t* pareada

Hipótesis nula: $H_0 : \mu_D = \Delta_0$

Estadístico de la prueba: $$T_0 = \frac{\overline{D} - \Delta_0}{S_D / \sqrt{n}} \tag{9-23}$$

Hipótesis alternativa	Región de rechazo
$H_1 : \mu_D \neq \Delta_0$	$t_0 > t_{\alpha/2,\, n-1}$ o $t_0 < -t_{\alpha/2,\, n-1}$
$H_1 : \mu_D > \Delta_0$	$t_0 > t_{\alpha,\, n-1}$
$H_1 : \mu_D < \Delta_0$	$t_0 < -t_{\alpha,\, n-1}$

En la ecuación 9-23, $\overline{D}$ es el promedio muestral de las n diferencias $D_1, D_2, ..., D_n$, y S_D es la desviación estándar muestral de estas diferencias.

EJEMPLO 9-9

En un artículo de *Journal of Strain Analysis* (Vol. 18, No. 2) se comparan varios métodos para predecir la resistencia al corte de vigas de placas de acero. En la tabla 9-4 se presentan los datos de dos de estos métodos, los procedimientos Karlsruhe y Lehigh, cuando se aplicaron a nueve vigas especificas. Quiere determinarse si hay alguna diferencia (en promedio) entre los dos métodos.

El procedimiento de ocho pasos se aplica como sigue:

1. El parámetro de interés es la diferencia de la resistencia al corte medio entre los dos métodos, por ejemplo, $\mu_D = \mu_1 - \mu_2 = 0$.
2. H_0: $\mu_D = 0$
3. H_1: $\mu_D \neq 0$
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$t_0 = \frac{\overline{d}}{s_D / \sqrt{n}}$$

6. Se rechaza H_0 si $t_0 > t_{0.025,\,8} = 2.306$ o si $t_0 < -t_{0.025,\,8} = -2.306$.
7. Cálculos: El promedio y la desviación estándar muestrales de las diferencias d_j son $\overline{d} = 0.2736$ y $s_D = 0.1356$, por lo que el estadístico de la prueba es

$$t_0 = \frac{\overline{d}}{s_D / \sqrt{n}} = \frac{0.2736}{0.1356 / \sqrt{9}} = 6.05$$

8. Conclusiones: Puesto que $t_0 = 6.05 > 2.306$, se concluye que los métodos de predicción de la resistencia al corte producen resultados diferentes. Específicamente, los datos indican que el método Karlsruhe produce, en promedio, predicciones de la resistencia al corte más altas que el método Lehigh. El valor P para $t_0 = 6.05$ es $P = 0.0002$, por lo que el estadístico de la prueba está bastante adentro de la región crítica.

Comparaciones pareadas y comparaciones no pareadas

Cuando se lleva a cabo un experimento comparativo, en ocasiones el investigador puede elegir entre el experimento pareado y el experimento de dos muestras (o no pareado). Si van a hacerse n mediciones en cada población, el estadístico de la prueba para dos muestras es

$$T_0 = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{S_p \sqrt{\frac{1}{n} + \frac{1}{n}}}$$

Tabla 9-4 Predicciones de la resistencia para nueve vigas de placas de acero (carga predicha/carga observada)

Viga	Método Karlsruhe	Método Lehigh	Diferencia d_j
S1/1	1.186	1.061	0.119
S2/1	1.151	0.992	0.159
S3/1	1.322	1.063	0.259
S4/1	1.339	1.062	0.277
S5/1	1.200	1.065	0.138
S2/1	1.402	1.178	0.224
S2/2	1.365	1.037	0.328
S2/3	1.537	1.086	0.451
S2/4	1.559	1.052	0.507

que se compararía con t_{2n-2}, y desde luego, el estadístico t pareado es

$$T_0 = \frac{\overline{D} - \Delta_0}{S_D / \sqrt{n}}$$

que se compara con t_{n-1}. Obsérvese que como

$$\overline{D} = \sum_{j=1}^{n} \frac{D_j}{n} = \sum_{j=1}^{n} \frac{(X_{1j} - X_{2j})}{n} = \sum_{j=1}^{n} \frac{X_{1j}}{n} - \sum_{j=1}^{n} \frac{X_{2j}}{n} = \overline{X}_1 - \overline{X}_2$$

los numeradores de ambos estadísticos son idénticos. Sin embargo, el denominador de la prueba t para dos muestras se basa en el supuesto de que X_1 y X_2 son *independientes.* En muchos experimentos pareados existe una fuerte correlación positiva ρ entre X_1 y X_2. Entonces puede demostrarse que

$$\begin{aligned} V(\overline{D}) &= V(\overline{X}_1 - \overline{X}_2 - \Delta_0) \\ &= V(\overline{X}_1) + V(\overline{X}_2) - 2\,\mathrm{cov}(\overline{X}_1, \overline{X}_2) \\ &= \frac{2\sigma^2(1-\rho)}{n} \end{aligned}$$

suponiendo que ambas poblaciones X_1 y X_2 tienen varianzas σ^2 idénticas. Además, S_D^2 / n estima la varianza de $\overline{D}$. Siempre que hay una correlación positiva dentro de los pares, el denominador de la prueba t pareada será menor que el denominador de la prueba t para dos muestras. Esto puede ocasionar que la prueba t para dos muestras produzca una subestimación considerable de la significación de los datos si se aplicara incorrectamente a muestras pareadas.

Aun cuando el pareado con frecuencia llevará a un valor menor de la varianza de $\overline{X}_1 - \overline{X}_2$, tiene una desventaja; es decir, la prueba t pareada lleva a una pérdida de $n - 1$ grados de libertad en comparación con la prueba t para dos muestras. En general, se sabe que al aumentar los grados de libertad de una prueba se incrementa la potencia contra cualquier valor alternativo fijo del parámetro.

Entonces, ¿cómo se decide cuál de las dos formas de conducir el experimento deberá usarse? ¿Las observaciones deberán parearse o no? Aun cuando no hay una respuesta general a esta pregunta, es posible ofrecer algunas pautas con base en el análisis anterior.

1. Si las unidades experimentales son relativamente homogéneas (σ pequeña) y la correlación dentro de los pares es pequeña, la ganancia en precisión atribuible al pareado será cancelada por la pérdida de grados de libertad, por lo que deberá usarse un experimento de muestras independientes.
2. Si las unidades experimentales son relativamente heterogéneas (σ grande) y hay una correlación positiva grande entre los pares, deberá usarse el experimento pareado. De manera típica, este caso se presenta cuando las unidades experimentales son las *mismas* para ambos tratamientos; como en el ejemplo 9-9, donde se usaron las mismas vigas para probar los dos métodos.

La aplicación de estas reglas aún requiere criterio, ya que nunca se conoce el valor exacto de σ y ρ. Además, si el número de grados de libertad es grande (40 o 50, por ejemplo), entonces la pérdida de $n - 1$ grados de libertad por el pareado puede no ser grave. Sin embargo, si el número de grados de libertad es pequeño (10 o 20, por ejemplo), entonces perder la mitad de ellos es potencialmente serio si no se compensa con una precisión incrementada por el pareado.

Intervalo de confianza para μ_D

Para construir el intervalo de confianza para $\mu_D = \mu_1 - \mu_2$, se observa que

$$T = \frac{\overline{D} - \mu_D}{S_D / \sqrt{n}}$$

sigue una distribución t con $n - 1$ grados de libertad. Entonces, puesto que

$$P\left(-t_{\alpha/2,\, n-1} \le T \le t_{\alpha/2,\, n-1}\right) = 1 - \alpha$$

puede sustituirse T en la expresión anterior y realizar los pasos necesarios para aislar $\mu_D = \mu_1 - \mu_2$ entre las desigualdades. Esto lleva al siguiente intervalo de confianza del $100(1 - \alpha)$ por ciento para $\mu_1 - \mu_2$.

Definición: Intervalo de confianza para $\mu_1 - \mu_2$ para observaciones pareadas

Si $\overline{d}$ y s_d son la media y la desviación estándar muestrales de la diferencia de n pares aleatorios de mediciones que tienen una distribución normal, entonces un intervalo de confianza del $100(1 - \alpha)$ por ciento para la diferencia de las medias $\mu_D = \mu_1 - \mu_2$ es

$$\overline{d} - t_{\alpha/2,\, n-1} s_D / \sqrt{n} \le \mu_D \le \overline{d} + t_{\alpha/2,\, n-1} s_D / \sqrt{n} \qquad (9\text{-}24)$$

donde $t_{\alpha/2,\, n-1}$ es el punto porcentual $\alpha/2$ superior de la distribución t con $n - 1$ grados de libertad.

Tabla 9-5 Tiempo en segundos para estacionar en paralelo dos automóviles

Sujeto	Automóvil 1 (x_{1j})	Automóvil 2 (x_{2j})	Diferencia (d_j)
1	37.0	17.8	19.2
2	25.8	20.2	5.6
3	16.2	16.8	–0.6
4	24.2	41.4	–17.2
5	22.0	21.4	0.6
6	33.4	38.4	–5.0
7	23.8	16.8	7.0
8	58.2	32.2	26.0
9	33.6	27.8	5.8
10	24.4	23.2	1.2
11	23.4	29.6	–6.2
12	21.2	20.6	0.6
13	36.2	32.2	4.0
14	29.8	53.8	–24.0

Este intervalo de confianza también es válido para el caso en que $\sigma_1^2 \neq \sigma_2^2$, ya que s_D^2 estima $\sigma_D^2 = V(X_1 - X_2)$. Asimismo, para muestras grandes (por ejemplo, $n \geq 30$ pares), el supuesto explícito de normalidad no es necesario debido al teorema del límite central.

EJEMPLO 9-10

La revista *Human Factors* reporta un estudio en el que se pidió a $n = 14$ sujetos estacionar en paralelo dos automóviles que tenían distancias entre ejes y radios de giro muy diferentes. Se registraron los tiempos en segundos para cada sujeto, los cuales se presentan en la tabla 9-5. A partir de la columna de las diferencias observadas se calculan $\bar{d} = 1.21$ y $s_d = 12.68$. Por la ecuación 9-24 se encuentra que el intervalo de confianza de 90% para $\mu_D = \mu_1 - \mu_2$ es:

$$\bar{d} - t_{0.05,13} s_D / \sqrt{n} \leq \mu_D \leq \bar{d} + t_{0.05,13} s_D / \sqrt{n}$$

$$1.21 - 1.771(12.68)/\sqrt{14} \leq \mu_D \leq 1.21 + 1.771(12.68)/\sqrt{14}$$

$$-4.79 \leq \mu_D \leq 7.21$$

Obsérvese que el intervalo de confianza para μ_D incluye a cero. Esto implica que, con el nivel de confianza de 90%, los datos no apoyan la afirmación de que los dos automóviles tienen tiempos de estacionamiento medios $\mu_1 - \mu_2$ diferentes. Es decir, el valor $\mu_D = \mu_1 - \mu_2 = 0$ no es inconsistente con los datos observados.

EJERCICIOS DE LA SECCIÓN 9-4

9-29. Considérese el experimento de la resistencia al corte que se describe en el ejemplo 9-9. Construya un intervalo de confianza de 95% para la diferencia de las medias de la resistencia al corte de los dos métodos. ¿El resultado que se obtiene es consistente con las conclusiones del ejemplo 9-9? Explique por qué.

9-30. Considérese de nuevo el experimento de la resistencia al corte que se describe en el ejemplo 9-9. ¿Es necesario que las dos resisten-

cias al corte individuales tengan una distribución normal para que la prueba t pareada sea apropiada, o es sólo la diferencia de las resistencias al corte la que debe ser normal? Use una gráfica de probabilidad normal para investigar el supuesto de normalidad.

9-31. Considérense los datos para estacionar los dos automóviles del ejemplo 9-10. Use la prueba t pareada para investigar la afirmación de que los dos tipos de automóvil tienen niveles diferentes de dificultad para estacionarlos en paralelo. Use $\alpha = 0.10$. Compare los resul tados con el intervalo de confianza construido en el ejemplo 9-10 y comente por qué son iguales o diferentes.

9-32. Considérense de nuevo los datos para estacionar los dos automóviles del ejemplo 9-10. Investigue el supuesto de que las diferencias en los tiempos de estacionamiento siguen una distribución normal.

9-33. El administrador de una flota de automóviles está probando dos marcas de llantas radiales. Instala una llanta de cada marca al azar en las dos ruedas traseras de ocho automóviles y los usa hasta que las llantas se desgastan. Los datos se presentan a continuación (en kilómetros). Encuentre un intervalo de confianza de 99% para la diferencia de la vida media. ¿Qué marca preferiría usted, con base en este cálculo?

Automóvil	Marca 1	Marca 2
1	36 925	34 318
2	45 300	42 280
3	36 240	35 500
4	32 100	31 950
5	37 210	38 015
6	48 360	47 800
7	38 200	37 810
8	33 500	33 215

9-34. Un especialista en computación está investigando la utilidad de dos lenguajes de diseño diferentes para mejorar las tareas de programación. Se pide a 12 programadores expertos, familiarizados con ambos lenguajes, que codifiquen una función estándar en ambos lenguajes, y se registra el tiempo en minutos. Los datos se presentan a continuación:

	Tiempo	
Programador	Lenguaje de diseño 1	Lenguaje de diseño 2
1	17	18
2	16	14
3	21	19
4	14	11
5	18	23
6	24	21
7	16	10
8	14	13
9	21	19
10	23	24
11	13	15
12	18	20

a) Encuentre un intervalo de confianza de 95% para la diferencia de las medias del tiempo de codificación. ¿Hay algún indicio de que uno de los lenguajes de diseño sea preferible?

b) ¿El supuesto de que la diferencia del tiempo de codificación tiene una distribución normal es razonable? Presente evidencias para apoyar su respuesta.

9-35. Quince adultos varones con edades entre 35 y 50 años participaron en un estudio para evaluar el efecto de la dieta y el ejercicio en los niveles de colesterol de la sangre. Se midió el colesterol total de cada sujeto al principio,

	Nivel de colesterol de la sangre	
Sujeto	Antes	Después
1	265	229
2	240	231
3	258	227
4	295	240
5	251	238
6	245	241
7	287	234
8	314	256
9	260	247
10	279	239
11	283	246
12	240	218
13	238	219
14	225	226
15	247	233

y tres meses después de participar en un programa de ejercicios aeróbicos y de cambiar a una dieta baja en grasas. Los datos se presentan en la tabla adjunta. ¿Los datos apoyan la afirmación de que la dieta baja en grasas y los ejercicios aeróbicos son de valor para producir una reducción en la media de los niveles de colesterol de la sangre? Use $\alpha = 0.05$.

9-36. Un artículo de *Journal of Aircraft* (Vol. 23, pp. 859, 864) describe la formulación de un nuevo método de análisis de placas equivalente que puede modelar estructuras aeronáuticas, tales como chumaceras acodadas de alas y que produce resultados similares al método de análisis de elementos finitos que requiere un número considerablemente mayor de cálculos. Se calcularon las frecuencias de vibración natural para la estructura de las chumaceras acodadas de alas usando ambos métodos, y los resultados de las siete primeras frecuencias naturales se presentan abajo.

Frecuencia	Elemento finito, ciclos/s	Placa equivalente, ciclos/s
1	14.58	14.76
2	48.52	49.10
3	97.22	99.99
4	113.99	117.53
5	174.73	181.22
6	212.72	220.14
7	277.38	294.80

a) ¿Los datos sugieren que los dos métodos producen el mismo valor medio de la frecuencia de vibración natural? Use $\alpha = 0.05$.

b) Encuentre un intervalo de confianza de 95% para la diferencia media entre los dos métodos.

9-37. Diez individuos participaron en un programa de modificación de su dieta para estimular la pérdida de peso. Sus pesos antes y después del programa se presentan en la tabla siguiente. ¿Hay evidencia en favor de la afirmación de que este programa particular para modificar la dieta es efectivo para producir una reducción en el peso medio? Use $\alpha = 0.05$.

Sujeto	Antes	Después
1	195	187
2	213	195
3	247	221
4	201	190
5	187	175
6	210	197
7	215	199
8	246	221
9	294	278
10	310	285

9-38. Pueden usarse dos pruebas analíticas diferentes para determinar el nivel de impureza en aleaciones de acero. Se prueban ocho muestras usando ambos procedimientos, cuyos resultados se presentan en la tabla siguiente. ¿Hay evidencia suficiente para concluir que ambas pruebas producen el mismo nivel de impureza medio usando $\alpha = 0.01$?

Muestra	Prueba 1	Prueba 2
1	1.2	1.4
2	1.3	1.7
3	1.5	1.5
4	1.4	1.3
5	1.7	2.0
6	1.8	2.1
7	1.4	1.7
8	1.3	1.6

9-39. Considérense los datos de la pérdida de peso del ejercicio 9-37. ¿Hay evidencia en favor de la afirmación de que este programa particular de modificación de la dieta dará como resultado una pérdida de peso promedio de al menos 10 libras? Use $\alpha = 0.05$.

9-40. Considérese el experimento de pérdida de peso del ejercicio 9-37. Suponga que, si el programa de modificación de la dieta da como resultado una pérdida de peso promedio de al menos 10 libras, es importante detectarlo con una probabilidad de al menos 0.90. ¿El uso de 10 sujetos fue un tamaño de la muestra adecuado? De no ser así, ¿cuántos sujetos se deberían haber usado?

9-5 INFERENCIA SOBRE LAS VARIANZAS DE DOS POBLACIONES NORMALES

Se introducen ahora las pruebas y los intervalos de confianza para las varianzas de las dos poblaciones ilustradas en la figura 9-1. Se supondrá que ambas poblaciones son normales. Tanto el procedimiento de prueba de hipótesis como la construcción del intervalo de confianza son relativamente sensibles al supuesto de normalidad.

9-5.1 Pruebas de hipótesis sobre el cociente de dos varianzas

Suponga que dos poblaciones normales independientes son de interés, donde las medias y las varianzas poblacionales, por ejemplo, μ_1, σ_1^2, μ_2 y σ_2^2, son desconocidas. Quieren probarse hipótesis acerca de la igualdad de las dos varianzas, por ejemplo, H_0: $\sigma_1^2 = \sigma_2^2$. Suponga que se cuenta con dos muestras aleatorias de tamaño n_1 de la población 1 y de tamaño n_2 de la población 2, y sean S_1^2 y S_2^2 las varianzas muestrales. Quieren probarse las hipótesis

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2 \tag{9-25}$$

El desarrollo de un procedimiento de prueba para estas hipótesis requiere una nueva distribución de probabilidad.

La distribución F

Una de las distribuciones más útiles de la estadística es la distribución F. La variable aleatoria F se define como el cociente de dos variables aleatorias ji-cuadrada independientes, dividida cada una por el número de sus grados de libertad. Es decir,

$$F = \frac{W/u}{Y/v} \tag{9-26}$$

donde W y Y son variables aleatorias ji-cuadrada independientes con u y v grados de libertad, respectivamente. Se enuncia ahora formalmente la distribución de muestreo de F.

La distribución F

Sean W y Y variables aleatorias ji-cuadrada independientes con u y v grados de libertad, respectivamente. Entonces el cociente

$$F = \frac{W/u}{Y/v} \tag{9-27}$$

tiene la función de densidad de probabilidad

$$f(x) = \frac{\Gamma\left(\frac{u+v}{2}\right)\left(\frac{u}{v}\right)^{u/2} x^{(u/2)-1}}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{v}{2}\right)\left[\left(\frac{u}{v}\right)x+1\right]^{(u+v)/2}} \quad 0 < x < \infty \tag{9-28}$$

y se dice que sigue una distribución F con u grados de libertad en el numerador y v grados de libertad en el denominador. Por lo general se abrevia $F_{u,v}$.

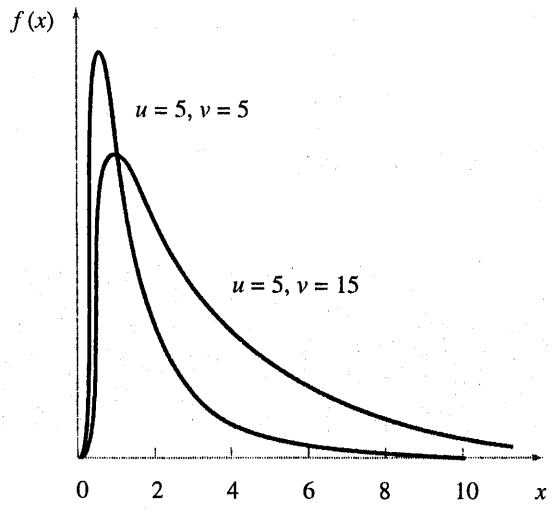


Figura 9-4 Funciones de densidad de probabilidad de dos distribuciones F.

La media y la varianza de la distribución F son $\mu = v/(v-2)$ para $v > 2$, y

$$\sigma^2 = \frac{2v^2(u+v-2)}{u(v-2)^2(v-4)}, \qquad v > 4$$

En la figura 9-4 se ilustran dos distribuciones F. La variable aleatoria F es no negativa, y la distribución está sesgada a la derecha. La distribución F se parece mucho a la distribución ji-cuadrada de la figura 8-15; sin embargo, los dos parámetros u y v proporcionan flexibilidad adicional en cuanto a la forma.

En la tabla V del apéndice se dan los puntos porcentuales de la distribución F. Sea $f_{\alpha, u, v}$ el punto porcentual de la distribución F, con u grados de libertad en el numerador y v grados de libertad en el denominador, tal que la probabilidad de que la variable aleatoria F exceda este valor es

$$P(F > f_{\alpha, u, v}) = \int_{f_{\alpha, u, v}}^{\infty} f(x)dx = \alpha$$

Esto se ilustra en la figura 9-5. Por ejemplo, si $u = 5$ y $v = 10$, por la tabla V del apéndice se encuentra que

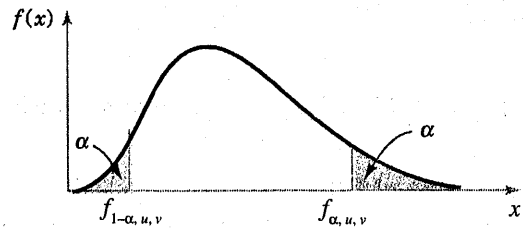


Figura 9-5 Puntos porcentuales superior e inferior de la distribución F.

$$P(F > f_{0.05,5,10}) = P(F_{5,10} > 3.33) = 0.05$$

Es decir, el punto porcentual 5 superior de $F_{5,10}$ es $f_{0.05,5,10} = 3.33$.

La tabla V sólo contiene puntos porcentuales de la cola superior (para valores seleccionados de $f_{\alpha,u,v}$ para $\alpha \leq 0.25$) de la distribución F. Los puntos porcentuales $f_{1-\alpha,u,v}$ de la cola inferior pueden encontrarse como sigue:

$$f_{1-\alpha,u,v} = \frac{1}{f_{\alpha,v,u}} \tag{9-29}$$

Por ejemplo, para encontrar el punto porcentual $f_{0.95,5,10}$ de la cola inferior se observa que

$$f_{0.95,5,10} = \frac{1}{f_{0.05,10,5}} = \frac{1}{4.74} = 0.211$$

El procedimiento de prueba

Un procedimiento de prueba de hipótesis para la igualdad de dos varianzas se basa en el siguiente resultado:

Sea $X_{11}, X_{12}, \ldots, X_{2n_1}$ una muestra aleatoria de una población normal con media μ_1 y varianza σ_1^2, y sea $X_{21}, X_{22}, \ldots, X_{2n_2}$ una muestra aleatoria de una segunda población normal con media μ_2 y varianza σ_2^2. Suponga que ambas poblaciones normales son independientes. Sean S_1^2 y S_2^2 las varianzas muestrales. Entonces el cociente

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

tiene una distribución F con $n_1 - 1$ grados de libertad en el numerador y $n_2 - 1$ grados de libertad en el denominador.

Este resultado se basa en el hecho de que $(n_1 - 1)S_1^2/\sigma_1^2$ es una variable aleatoria ji-cuadrada con $n_1 - 1$ grados de libertad, que $(n_2 - 1)S_2^2/\sigma_2^2$ es una variable aleatoria ji-cuadrada con $n_2 - 1$ grados de libertad, y que las dos poblaciones normales son independientes. Evidentemente, bajo la hipótesis nula $H_0 : \sigma_1^2 = \sigma_2^2$, el cociente $F_0 = S_1^2/S_2^2$ tiene una distribución F_{n_1-1,n_2-1}. Esta es la base del procedimiento de prueba siguiente:

Hipótesis nula: $H_0 : \sigma_1^2 = \sigma_2^2$

Estadístico de la prueba: $F_0 = \dfrac{S_1^2}{S_2^2}$ (9-30)

Hipótesis alternativas	Criterio de rechazo	
$H_1 : \sigma_1^2 \neq \sigma_2^2$	$f_0 > f_{\alpha/2, n_1-1, n_2-1}$	$f_0 < f_{1-\alpha/2, n_1-1, n_2-1}$
$H_1 : \sigma_1^2 > \sigma_2^2$	$f_0 > f_{\alpha, n_1-1, n_2-1}$	
$H_1 : \sigma_1^2 < \sigma_2^2$	$f_0 < f_{1-\alpha, n_1-1, n_2-1}$	

EJEMPLO 9-11

Las capas de óxido sobre obleas semiconductoras se someten a grabado en una mezcla de gases para conseguir el espesor correcto. La variablidad en el espesor de estas capas de óxido es una característica crítica de las obleas, y es deseable una variabilidad reducida en las fases subsecuentes del procesamiento. Se estudian dos mezclas diferentes de gases a fin de determinar si una de ellas es mejor para reducir la variabilidad del espesor del óxido. Se someten a corrosión 20 obleas en cada gas. Las desviaciones estándar muestrales del espesor del óxido son $s_1 = 1.96$ angstroms y $s_2 = 2.13$ angstroms, respectivamente. ¿Hay evidencia que indique que deba preferirse alguno de los dos gases? Use $\alpha = 0.05$.

El procedimiento de prueba de hipótesis de ocho pasos puede aplicarse en este problema como sigue:

1. Los parámetros de interés son las varianzas del espesor del óxido, σ_1^2 y σ_2^2. Se supondrá que el espesor del óxido es una variable aleatoria normal para ambas mezclas de gas.
2. $H_0 : \sigma_1^2 = \sigma_2^2$
3. $H_1 : \sigma_1^2 \neq \sigma_2^2$
4. $\alpha = 0.05$
5. El estadístico de la prueba está dado por la ecuación 9-30:

$$f_0 = \frac{s_1^2}{s_2^2}$$

6. Puesto que $n_1 = n_2 = 20$, se rechazará $H_0 : \sigma_1^2 = \sigma_2^2$ si $f_0 > f_{0.025, 19, 19} = 2.53$ o si $f_0 < f_{0.975, 19, 19} = 1/f_{0.025, 19, 19} = 1/2.53 = 0.40$.
7. Cálculos: Puesto que $s_1^2 = (1.96)^2 = 3.84$ y $s_2^2 = (2.13)^2 = 4.54$, el estadístico de la prueba es

$$f_0 = \frac{s_1^2}{s_2^2} = \frac{3.84}{4.54} = 0.85$$

8. Conclusiones: Puesto que $f_{0.975,\,19,\,19} = 0.40 < f_0 = 0.85 < f_{0.025,\,19,\,19} = 2.53$, no puede rechazarse la hipótesis nula $H_0 : \sigma_1^2 = \sigma_2^2$ con el nivel de significación 0.05. Por lo tanto, no hay evidencia sólida que indique que alguno de los dos gases dé como resultado una varianza menor del espesor del óxido.

También es posible encontrar un valor P para el estadístico F del ejemplo 9-11. Puesto que $f_{0.50,\,19,\,19} = 1.00$, el valor calculado del estadístico de la prueba $f_0 = s_1^2 / s_2^2 = 3.84/4.54 = 0.85$ se encuentra más cerca de la cola inferior de la distribución F que de la cola superior. La probabilidad de que una variable aleatoria F con 19 grados de libertad en el numerador y el denominador sea menor que 0.85 es 0.3634. Puesto que es arbitrario cuál de las poblaciones se identifica como "uno", el estadístico de la prueba pudo haberse calculado como $f_0 = 4.54/3.84 = 1.18$. La probabilidad de que una variable aleatoria F con 19 grados de libertad en el numerador y el denominador exceda 1.18 es 0.3610. Por lo tanto, el valor P para el estadístico de la prueba $f_0 = 0.85$ es la suma de estas dos probabilidades, o bien $P = 0.3634 + 0.3610 = 0.7244$. Puesto que el valor P excede 0.05, no puede rechazarse la hipótesis nula $H_0 : \sigma_1^2 = \sigma_2^2$. (Las probabilidades dadas arriba se calcularon con una calculadora manual.)

9-5.2 Error β y elección del tamaño de la muestra

En las gráficas VI*o*, VI*p*, VI*q* y VI*r* del apéndice se dan las curvas de operación característica de la prueba F que se presentó en la sección 9-5.1 para $\alpha = 0.05$ y $\alpha = 0.01$, suponiendo que $n_1 = n_2 = n$. Las gráficas VI*o* y VI*p* se usan con las hipótesis alternativas de dos colas. En ellas se grafica β contra el parámetro de las abscisas

$$\lambda = \frac{\sigma_1}{\sigma_2} \tag{9-31}$$

para diferentes valores $n_1 = n_2 = n$. Las gráficas VI*q* y VI*r* se usan para las hipótesis alternativas de una cola.

EJEMPLO 9-12

En el problema del grabado con óxido de la oblea semiconductora del ejemplo 9-11, suponga que uno de los gases dio como resultado una desviación estándar del espesor del óxido que es la mitad de la desviación estándar del espesor del óxido del otro gas. Si quisiera detectarse esta situación con una probabilidad de al menos 0.80, ¿el tamaño de la muestra $n_1 = n_2 = 20$ es adecuado?

Obsérvese que si una de las desviaciones estándar es la mitad de la otra, entonces

$$\lambda = \frac{\sigma_1}{\sigma_2} = 2$$

En la gráfica VI*o* del apéndice, con $n_1 = n_2 = n = 20$ y $\lambda = 2$, se encuentra que $\beta \cong 0.20$. Por lo tanto, si $\beta = 0.20$, la potencia del prueba (que es la probabilidad de que la diferencia de las desviaciones estándar se detectará en la prueba) es 0.80, y se concluye que los tamaños de las muestras $n_1 = n_2 = 20$ son adecuados.

9-5.3 Intervalo de confianza sobre el cociente de dos varianzas

Para encontrar el intervalo de confianza para σ_1^2/σ_2^2, recuérdese que la distribución de muestreo de

$$F = \frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}$$

es una distribución F con $n_2 - 1$ y $n_1 - 1$ grados de libertad. Por lo tanto,

$$P\left(f_{1-\alpha/2,n_2-1,n_1-1} \le F \le f_{\alpha/2,n_2-1,n_1-1}\right) = 1-\alpha$$

Al sustituir F y manipular las desigualdades se llega al intervalo de confianza del $100(1-\alpha)$ por ciento para σ_1^2/σ_2^2.

Definición: Intervalo de confianza para el cociente de las varianzas de dos distribuciones normales

Si s_1^2 y s_2^2 son las varianzas muestrales de las muestras aleatorias de tamaños n_1 y n_2, respectivamente, de dos poblaciones normales independientes con varianzas σ_1^2 y σ_2^2 desconocidas, entonces un intervalo de confianza de $100(1-\alpha)$ por ciento para el cociente σ_1^2/σ_2^2 es

$$\frac{s_1^2}{s_2^2} f_{1-\alpha/2,n_2-1,n_1-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} f_{\alpha/2,n_2-1,n_1-1} \qquad (9\text{-}32)$$

donde $f_{\alpha/2,n_2-1,n_1-1}$ y $f_{1-\alpha/2,n_2-1,n_1-1}$ son los puntos porcentuales $\alpha/2$ superior e inferior de la distribución F con $n_2 - 1$ y $n_1 - 1$ grados de libertad en el numerador y el denominador, respectivamente.

EJEMPLO 9-13

Una compañía fabrica propulsores que se usan en turbinas de aviones. Una de las operaciones implica pulir un terminado superficial particular en un componente de una aleación de titanio. Pueden usarse dos procesos de pulido diferentes, y ambos procesos pueden producir piezas con una aspereza superficial media idéntica. Al ingeniero de manufactura le gustaría seleccionar el proceso que presente la variabilidad menor en la aspereza superficial. Una muestra aleatoria de $n_1 = 11$ piezas del primer proceso da como resultado una desviación estándar muestral $s_1 = 5.1$ micropulgadas, y una muestra aleatoria de $n_2 = 16$ piezas del segundo proceso da como resultado una desviación estándar muestral de $s_2 = 4.7$ micropulgadas. Se encontrará un intervalo de confianza de 90% para el cociente de las dos varianzas σ_1^2/σ_2^2.

Suponiendo que los dos procesos son independientes y que la aspereza superficial tiene una distribución normal, puede usarse la ecuación 9-32 como sigue:

$$\frac{s_1^2}{s_2^2} f_{0.95,15,10} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} f_{0.05,15,10}$$

$$\frac{(5.1)^2}{(4.7)^2} 0.39 \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{(5.1)^2}{(4.7)^2} 2.85$$

o bien

$$0.46 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3.56$$

Obsérvese que se ha usado la ecuación 9-30 para encontrar $f_{0.95,\,15,\,10} = 1/f_{0.05,\,10,\,15} = 1/2.54 = 0.39$. Puesto que este intervalo de confianza incluye la unidad, no puede afirmarse que las desviaciones estándar de la aspereza superficial de los dos procesos son diferentes con el nivel de confianza de 90%.

EJERCICIOS DE LA SECCIÓN 9-5

9-41. Para una distribución F, encuentre lo siguiente:

a) $f_{0.25,\,5,\,10}$ d) $f_{0.75,\,5,\,10}$
b) $f_{0.10,\,24,\,9}$ e) $f_{0.90,\,24,\,9}$
c) $f_{0.05,\,8,\,15}$ f) $f_{0.95,\,8,\,15}$

9-42. Para una distribución F, encuentre lo siguiente:

a) $f_{0.25,\,7,\,15}$ d) $f_{0.75,\,7,\,15}$
b) $f_{0.10,\,10,\,12}$ e) $f_{0.90,\,10,\,12}$
c) $f_{0.01,\,20,\,10}$ f) $f_{0.99,\,20,\,10}$

9-43. Dos compañías químicas pueden surtir una materia prima. La concentración de un elemento particular en este material es importante. La concentración media de ambos proveedores es la misma, pero se sospecha que la variabilidad de la concentración puede diferir entre las dos compañías. La desviación estándar de la concentración en una muestra aleatoria de $n_1 = 10$ lotes producidos por la compañía 1 es $s_1 = 4.7$ g/l, mientras que para la compañía 2 una muestra aleatoria de $n_2 = 16$ lotes da como resultado $s_2 = 5.8$ g/l. ¿Hay evidencia suficiente para concluir que las dos varianzas poblacionales difieren? Use $\alpha = 0.05$.

9-44. Considérense los datos de la rapidez de grabado del ejercicio 9-19. Pruebe la hipótesis $H_0 : \sigma_1^2 = \sigma_2^2$ contra $H_1 : \sigma_1^2 \neq \sigma_2^2$ usando $\alpha = 0.05$, y saque conclusiones.

9-45. Considérense los datos de la rapidez de grabado del ejercicio 9-19. Suponga que si la varianza de una de las poblaciones es el doble de la otra, quiere detectarse esta situación con una probabilidad de al menos 0.90 (usando $\alpha = 0.05$). ¿Los tamaños de las muestras $n_1 = n_2 = 0$ son adecuados?

9-46. Considérense los datos de los diámetros de ejercicio 9-15. Construya lo siguiente:

a) Un intervalo de confianza de dos colas de 90% para σ_1/σ_2.
b) Un intervalo de confianza de dos colas de 95% para σ_1/σ_2. Comente la comparación de la anchura de este intervalo con la del intervalo del inciso a).
c) Un intervalo de confianza inferior de 90% para σ_1/σ_2.

9-47. Considérense los datos de la espuma del ejercicio 9-16. Construya lo siguiente:

a) Un intervalo de confianza de dos colas de 90% para σ_1^2/σ_2^2.
b) Un intervalo de confianza de dos colas de 95% para σ_1^2/σ_2^2. Comente la comparación de la anchura de este intervalo con la del intervalo del inciso a).
c) Un intervalo de confianza inferior de 90% para σ_1/σ_2.

9-48. Considérense los datos de la rapidez de la película del ejercicio 9-22. Pruebe H_0: $\sigma_1^2 = \sigma_2^2$ contra $H_1 : \sigma_1^2 \neq \sigma_2^2$. Use $\alpha = 0.02$.

9-49. Considérense los datos de la resistencia a impactos del engrane del ejercicio 9-20. ¿Hay

evidencia suficiente para concluir que la varianza de la resistencia a impactos de los dos proveedores es diferente? Use $\alpha = 0.05$.

9-50. Considérense los datos del punto de fusión del ejercicio 9-23. ¿Los datos muestrales apoyan la afirmación de que ambas aleaciones tienen la misma varianza del punto de fusión? Use $\alpha = 0.05$ y llegue a una conclusión.

9-51. En el ejercicio 9-26 se presentaron las mediciones del espesor de un recubrimiento plástico con dos temperaturas de aplicación diferentes. Pruebe $H_0: \sigma_1^2 = \sigma_2^2$ contra H_1: $\sigma_1^2 \neq \sigma_2^2$ usando $\alpha = 0.01$.

9-52. Se realizó un estudio a fin de determinar si los hombres y las mujeres difieren en su repetibilidad en el ensamblaje de componentes en tarjetas de circuitos impresos. Se seleccionaron dos muestras de 25 hombres y 21 mujeres, y cada sujeto ensambló las unidades. Las dos desviaciones estándar muestrales del tiempo de ensamblaje fueron $s_{\text{hombres}} = 0.98$ min y $s_{\text{mujeres}} = 1.02$ min. ¿Hay evidencia en favor de la afirmación de que los hombres y las mujeres difieren en su repetibilidad en esta tarea de ensamblaje? Use $\alpha = 0.02$ y enuncie cualquier supuesto necesario acerca de la distribución fundamental de los datos.

9-53. Considérese de nuevo el experimento de la repetibilidad del ensamblaje descrito en el ejercicio 9-52. Encuentre un intervalo de confianza de 98% para el cociente de las dos varianzas. Proponga una interpretación del intervalo.

9-54. Considérese de nuevo el experimento de la rapidez de la película del ejercicio 9-22. Suponga que la desviación estándar de una de las poblaciones es 50% mayor que la otra. ¿El tamaño de la muestra $n_1 = n_2 = 8$ es adecuado para detectar esta diferencia con una alta probabilidad? Use $\alpha = 0.01$ para responder esta pregunta.

9-6 INFERENCIA SOBRE PROPORCIONES DE DOS POBLACIONES

Se considera ahora el caso en que hay dos parámetros binomiales de interés, por ejemplo, p_1 y p_2, y quieren hacerse inferencias acerca de estas proporciones. Se presentarán procedimientos de prueba de hipótesis de muestras grandes e intervalos de confianza basados en la aproximación normal de la distribución binomial.

9-6.1 Prueba de una muestra grande para H_0: $p_1 = p_2$

Suponga que las dos muestras aleatorias independientes de tamaños n_1 y n_2 se toman de dos poblaciones, y sea que X_1 y X_2 representen el número de observaciones que pertenecen a la clase de interés en las muestras 1 y 2, respectivamente. Suponga asimismo que se aplica la aproximación normal de la distribución binomial a cada una de las poblaciones, de tal modo que los estimadores de las proporciones poblacionales $\hat{P}_1 = X_1/n_1$ y $\hat{P}_2 = X_2/n_2$ tienen distribuciones normales aproximadas. El interés se encuentra en probar las hipótesis

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2$$

El estadístico

$$Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \tag{9-33}$$

tiene una distribución aproximadamente normal estándar y es la base de una prueba de H_0: $p_1 = p_2$. Específicamente, si la hipótesis nula H_0: $p_1 = p_2$ es verdadera, entonces al usar el hecho de que $p_1 = p_2 = p$, la variable aleatoria

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{p(1-p)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

tiene una distribución aproximada $N(0, 1)$. Un estimador del parámetro común p es

$$\hat{P} = \frac{X_1 + X_2}{n_1 + n_2}$$

El estadístico de la prueba para H_0: $p_1 = p_2$ es entonces

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

Esto lleva a los procedimientos de prueba que se describen a continuación:

Hipótesis nula: H_0: $p_1 = p_2$

Estadístico de la prueba:

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \tag{9-34}$$

Hipótesis alternativas	Criterio de rechazo
H_1: $p_1 \neq p_2$	$z_0 > z_{\alpha/2}$ o $z_0 < -z_{\alpha/2}$
H_1: $p_1 > p_2$	$z_0 > z_\alpha$
H_1: $p_1 < p_2$	$z_0 < -z_\alpha$

EJEMPLO 9-14

Se evalúan dos tipos diferentes de soluciones de pulido para su posible uso en una operación de pulido rotatorio en la fabricación de lentes intraoculares usados en el ojo humano después de una cirugía de cataratas. Se pulieron rotatoriamente 300 lentes usando la primera solución de pulido, de los cuales 253 no tuvieron defectos producidos por el pulido. Otros 300 lentes se pulieron rotatoriamente usando la segunda solución de pulido, y 196 lentes fueron satisfactorios al término de la operación. ¿Hay alguna razón para creer que las dos soluciones de pulido difieren? Use $\alpha = 0.01$.

El procedimiento de prueba de hipótesis de ocho pasos lleva a los siguientes resultados:

1. Los parámetros de interés son p_1 y p_2, la proporción de lentes que son satisfactorios después del pulido rotatorio con los fluidos de pulido 1 o 2.
2. H_0: $p_1 = p_2$
3. H_1: $p_1 \neq p_2$
4. $\alpha = 0.01$
5. El estadístico de la prueba es

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

donde $\hat{p}_1 = 253/300 = 0.8433$, $\hat{p}_2 = 196/300 = 0.6533$, $n_1 = n_2 = 300$, y

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{253 + 196}{300 + 300} = 0.7483$$

6. Se rechaza H_0: $p_1 = p_2$ si $z_0 > z_{0.005} = 2.58$ o si $z_0 < -z_{0.005} = -2.58$.
7. Cálculos: El valor del estadístico de la prueba es

$$z_0 = \frac{0.8433 - 0.6533}{\sqrt{0.7483(0.2517)\left(\frac{1}{300} + \frac{1}{300}\right)}} = 5.36$$

8. Conclusiones: Puesto que $z_0 = 5.36 > z_{0.005} = 2.58$, se rechaza la hipótesis nula. Obsérvese que el valor P es $P \simeq 8.32 \times 10^{-8}$. Hay evidencia sólida en favor de la afirmación de que los dos fluidos de pulido son diferentes. El fluido 1 produce una proporción mayor de lentes no defectuosos.

9-6.2 Error β y elección del tamaño de la muestra

El cálculo del error β para la prueba anterior es un tanto más complicado que en el caso de una sola muestra. El problema es que el denominador de Z_0 es una estimación de la desviación

estándar de $\hat{P}_1 - \hat{P}_2$ bajo el supuesto de que $p_1 = p_2 = p$. Cuando H_0: $p_1 = p_2$ es falsa, la desviación estándar de $\hat{P}_1 - \hat{P}_2$ es

$$\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \tag{9-35}$$

Si la hipótesis alternativa es de dos colas, el error β es

$$\beta = \Phi\left(\frac{z_{\alpha/2}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}}\right) - \Phi\left(\frac{-z_{\alpha/2}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}}\right) \tag{9-36}$$

donde

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$\bar{q} = \frac{n_1(1-p_1) + n_2(1-p_2)}{n_1 + n_2}$$

y $\sigma_{\hat{P}_1 - \hat{P}_2}$ está dada por la ecuación 9-35.

Si la hipótesis alternativa es H_1: $p_1 > p_2$, entonces

$$\beta = \Phi\left(\frac{z_{\alpha}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}}\right) \tag{9-37}$$

y si la hipótesis alternativa es H_1: $p_1 < p_2$, entonces

$$\beta = 1 - \Phi\left(\frac{-z_{\alpha}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}}\right) \tag{9-38}$$

Para un par especificado de valores p_1 y p_2 pueden encontrarse los tamaños de las muestras $n_1 = n_2 = n$ necesarios para dar la prueba de tamaño α que tiene un error β tipo II especificado.

Para la hipótesis alternativa de dos colas, el tamaño de la muestra común es

$$n = \frac{\left(z_{\alpha/2}\sqrt{(p_1+p_2)(q_1+q_2)/2} + z_\beta\sqrt{p_1q_1+p_2q_2}\right)^2}{(p_1-p_2)^2} \tag{9-39}$$

donde $q_1 = 1 - p_1$ y $q_2 = 1 - p_2$.

Para una hipótesis alternativa de una cola, se reemplaza $z_{\alpha/2}$ en la ecuación 9-39 con z_α.

9-6.3 Intervalo de confianza para $p_1 - p_2$

El intervalo de confianza para $p_1 - p_2$ puede encontrarse directamente, pues se sabe que

$$Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}}$$

es una variable aleatoria normal estándar. Por tanto,

$$P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) \simeq 1 - \alpha$$

por lo que puede sustituirse Z en esta última expresión y usar un enfoque similar al que se empleó anteriormente para encontrar un intervalo de confianza del $100(1 - \alpha)$ por ciento aproximado para $p_1 - p_2$.

Definición: Intervalo de confianza para la diferencia de dos proporciones

Si $\hat{p}_1$ y $\hat{p}_2$ son las proporciones muestrales de las observaciones en dos muestras aleatorias independientes de tamaños n_1 y n_2 que pertenecen a la clase de interés, entonces un intervalo de confianza del $100(1 - \alpha)$ por ciento aproximado para la diferencia de las verdaderas proporciones $p_1 - p_2$ es

$$\hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$
$$\le p_1 - p_2 \le \hat{p}_1 - \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \tag{9-40}$$

donde $z_{\alpha/2}$ es el punto porcentual $\alpha/2$ superior de la distribución normal estándar.

EJEMPLO 9-15

Considérese el proceso de manufactura de los rodamientos para cigüeñales del ejemplo 8-16. Suponga que se introduce una modificación en el proceso de terminado superficial y que, después, se obtiene una segunda muestra aleatoria de 85 árboles. El número de árboles defectuosos en la segunda muestra es 8. Por lo tanto, puesto que $n_1 = 85$, $\hat{p}_1 = 0.12$, $n_2 = 85$ y $\hat{p}_2 = 8/85 = 0.09$, utilizando la ecuación 9-40 puede obtenerse un intervalo de confianza aproximado de 95% para la diferencia de la proporción de rodamientos defectuosos producidos con los dos procesos como sigue:

$$\hat{p}_1 - \hat{p}_2 - z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$
$$\leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

o bien

$$0.12 - 0.09 - 1.96\sqrt{\frac{0.12(0.88)}{85} + \frac{0.09(0.91)}{85}}$$
$$\leq p_1 - p_2 \leq 0.12 - 0.09 + 1.96\sqrt{\frac{0.12(0.88)}{85} + \frac{0.09(0.91)}{85}}$$

Al simplificar esta expresión se obtiene

$$-0.06 \leq p_1 - p_2 \leq 0.12$$

Este intervalo de confianza incluye a cero, por lo que, con base en los datos muestrales, parece improbable que las modificaciones hechas en el proceso del terminado superficial hayan reducido la proporción de rodamientos defectuosos que se producen.

EJERCICIOS DE LA SECCIÓN 9-6

9-55. Se usan dos tipos diferentes de máquinas de moldeado por inyección para hacer piezas de plástico. Una pieza se considera defectuosa si presenta una merma excesiva o está decolorada. Se seleccionan dos muestras aleatorias, cada una de tamaño 300, y se encuentran 15 piezas defectuosas en la muestra de la máquina 1, mientras que se encuentran 8 piezas defectuosas en la muestra de la máquina 2. ¿Es razonable concluir que ambas máquinas producen la misma proporción de piezas defectuosas usando $\alpha = 0.05$? Encuentre el valor P para esta prueba.

9-56. Considérese la situación descrita en el ejercicio 9-55. Suponga que $p_1 = 0.05$ y $p_2 = 0.01$.

a) Con los tamaños de las muestras dados aquí, ¿cuál es la potencia de la prueba para esta hipótesis alternativa de dos colas?

b) Determine el tamaño de la muestra que se necesita para detectar esta diferencia con una probabilidad de al menos 0.9. Use $\alpha = 0.05$.

9-57. Considérese la situación descrita en el ejercicio 9-55. Suponga que $p_1 = 0.05$ y $p_2 = 0.02$.

a) Con los tamaños de las muestras dados aquí, ¿cuál es la potencia de la prueba para esta hipótesis alternativa de dos colas?

b) Determine el tamaño de la muestra que se necesita para detectar esta diferencia con una probabilidad de al menos 0.9. Use $\alpha = 0.05$.

9-58. En una muestra aleatoria de 500 adultos residentes del condado Maricopa se encontró que 385 estaban en favor de aumentar el límite de velocidad a 75 mph, mientras que en otra muestra de 400 adultos residentes del condado Pima se encontró que 267 estaban en favor de aumentar el límite de velocidad. ¿Estos datos indican que hay una diferencia en el apoyo para aumentar el límite de velocidad entre los residentes de los dos condados? Use $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

9-59. Construya un intervalo de confianza de 95% para la diferencia de las dos proporciones de piezas defectuosas del ejercicio 9-55.

9-60. Construya un intervalo de confianza de 95% para la diferencia de las dos proporciones del ejercicio 9-58. Proponga una interpretación práctica de este intervalo.

9-7 TABLA DE RESUMEN PARA PROCEDIMIENTOS DE INFERENCIA PARA DOS MUESTRAS

En la tabla de las contrapastas del libro se resumen todos los procedimientos de inferencia para dos muestras presentados en este capítulo. La tabla contiene el postulado de la hipótesis nula, el estadístico de la prueba, los criterios de rechazo de las diferentes hipótesis alternativas, y las fórmulas para construir intervalos de confianza del $100(1 - \alpha)$ por ciento.

Ejercicios complementarios

9-61. Un especialista de almacén compró 25 resistores al vendedor 1 y 35 resistores al vendedor 2. Se mide la resistencia de cada resistor con los siguientes resultados:

Vendedor 1

96.8	100.0	100.3	98.5	98.3	98.2
99.6	99.4	99.9	101.1	103.7	97.7
99.7	101.1	97.7	98.6	101.9	101.0
99.4	99.8	99.1	99.6	101.2	98.2
98.6					

Vendedor 2

106.8	106.8	104.7	104.7	108.0	102.2
103.2	103.7	106.8	105.1	104.0	106.2
102.6	100.3	104.0	107.0	104.3	105.8
104.0	106.3	102.2	102.8	104.2	103.4
104.6	103.5	106.3	109.2	107.2	105.4
106.4	106.8	104.1	107.1	107.7	

a) ¿Qué supuesto sobre la distribución se necesita para probar la afirmación de que la varianza de la resistencia del producto del vendedor 1 no es significativamente diferente de la varianza de la resistencia del producto del vendedor 2? Lleve a cabo un procedimiento gráfico para verificar este supuesto.

b) Aplique un procedimiento de prueba de hipótesis estadística adecuado para determinar si el especialista de almacén puede afirmar que la varianza de la resistencia del producto del vendedor 1 difiere significativamente de la varianza de la resistencia del producto del vendedor 2.

9-62. Un artículo del *Journal of Materials Engineering* (Vol. II, No. 4, pp. 275-282) reportó los resultados de un experimento para determinar los mecanismos de falla de los recubri-

mientos de protección térmica por rocío de plasma. El esfuerzo de falla de un recubrimiento particular (NiCrAlZr) bajo dos condiciones de prueba diferentes es el siguiente:

Esfuerzo de falla ($\times 10^6$ Pa) después de nueve ciclos de 1 hr: 19.8, 18.5, 17.6, 16.7, 16.7, 14.8, 15.4, 14.1, 13.6

Esfuerzo de falla ($\times 10^6$ Pa) después de seis ciclos de 1 hr: 14.9, 12.7, 11.9, 11.4, 10.1, 7.9

a) ¿Qué supuestos se necesitan para construir intervalos de confianza para la diferencia del esfuerzo de falla medio bajo dos condiciones de prueba diferentes? Use gráficas de probabilidad normal de los datos para verificar estos supuestos.
b) Encuentre un intervalo de confianza de 99% para la diferencia del esfuerzo de falla medio bajo las dos condiciones de prueba diferentes.
c) Utilizando el intervalo de confianza construido en el inciso b), ¿la evidencia apoya la afirmación de que las primeras condiciones de prueba producen resultados más altos, en promedio, que las segundas? Explique su respuesta.

9-63. Considérese el ejercicio complementario 9-62.
a) Construya un intervalo de confianza de 95% para el cociente de las varianzas, s_1/s_2, del esfuerzo para la falla bajo las dos condiciones de prueba diferentes.
b) Use la respuesta del inciso b) para determinar si hay una diferencia significativa en las varianzas de las dos condiciones de prueba diferentes. Explique su respuesta.

9-64. Un producto dietético líquido anuncia en su publicidad que el uso del producto durante un mes da como resultado una pérdida de peso promedio de al menos 3 libras. Ocho sujetos usan el producto durante un mes y los datos de la pérdida de peso resultantes se presentan a continuación. Use procedimientos de prueba de hipótesis para responder las preguntas siguientes:

Sujeto	Peso inicial (lb)	Peso final (lb)
1	165	161
2	201	195
3	195	192
4	198	193
5	155	150
6	143	141
7	150	146
8	187	183

a) ¿Los datos apoyan la afirmación del fabricante del producto dietético con la probabilidad de un error tipo I fijada en 0.05?
b) ¿Los datos apoyan la afirmación del fabricante del producto dietético con la probabilidad de un error tipo I fijada en 0.01?
c) En un esfuerzo para mejorar las ventas, el fabricante está considerando cambiar la afirmación "al menos 3 libras" por "al menos 5 libras". Repita los incisos a) y b) para probar esta nueva afirmación.

9-65. Se investiga la resistencia a la ruptura del hilado de dos fabricantes. Se sabe por experiencia con los procesos de los fabricantes que $\sigma_1 = 5$ psi y $\sigma_2 = 4$ psi. Una muestra aleatoria de 20 fibras de prueba de cada fabricante da como resultado $\bar{x}_1 = 88$ psi y $\bar{x}_2 = 91$ psi, respectivamente.
a) Usando un intervalo de confianza de 90% para la diferencia de la resistencia a la ruptura media, comente acerca de si hay o no evidencia en favor de la afirmación de que el fabricante 2 produce hilado con una resistencia a la ruptura media más alta.
b) Usando un intervalo de confianza de 98% para la diferencia de la resistencia a la ruptura media, comente acerca de si hay o no evidencia en favor de la afirmación de que el fabricante 2 produce hilado con una resistencia a la ruptura media más alta.
c) Comente por qué los resultados de los incisos a) y b) son diferentes o iguales. ¿Cuál de los dos elegiría el lector para tomar una decisión y por qué?

9-66. Considérese el ejercicio complementario 9-55. Suponga que antes de colectar los datos se decide que el error al estimar $\mu_1 - \mu_2$ con $x_1 - x_2$ sea menor que 1.5 psi. Especifique el tamaño de la muestra para el siguiente porcentaje de confianza:

a) 90%

b) 98%

c) Comente el efecto de aumentar el porcentaje de confianza sobre el tamaño de la muestra que se necesita.

d) Repita los incisos a) al c) con un error de menos de 0.75 psi en lugar de 1.5 psi.

e) Comente el efecto de disminuir el error sobre el tamaño de la muestra que se necesita.

9-67. El experimento de la vacuna Salk contra la poliomielitis realizado en 1954 se enfocó en la efectividad de la vacuna para combatir la poliomielitis paralítica. Debido a que se pensó que sin un grupo de niños control no se contaría con bases sólidas para evaluar la eficacia de la vacuna Salk, se le administró la vacuna a un grupo, mientras que a un segundo grupo se le administró un placebo (visualmente idéntico a la vacuna pero sin efecto alguno). Por razones éticas, y debido a que se temía que el conocimiento de la administración de la vacuna afectaría los diagnósticos subsecuentes, el experimento se realizó en la modalidad de doble desconocimiento. Es decir, ni los sujetos ni las personas que las administraban sabían quién recibía la vacuna y quién recibía el placebo. Los datos reales de este experimento fueron los siguientes:

Grupo del placebo: n = 201 299: 110 casos de poliomielitis observados

Grupo de la vacuna: n = 200 745: 33 casos de poliomielitis observados

a) Use un procedimiento de prueba de hipótesis para determinar si la proporción de los niños que contrajeron poliomielitis de los dos grupos es estadísticamente diferente. Use una probabilidad del error tipo I igual a 0.05.

b) Repita el inciso a) usando una probabilidad del error tipo I igual a 0.01.

c) Compare las conclusiones de los incisos a) y b) y explique por qué son iguales o diferentes.

9-68. En una muestra aleatoria de 1 500 teléfonos domésticos de Phoenix en 1990, se encontró que 387 de los números no estaban en el directorio telefónico. En una muestra aleatoria de 1 200 teléfonos tomada el mismo año en Scottsdale se encontró que 310 no aparecían en el directorio.

a) Encuentre un intervalo de confianza de 95% para la diferencia de las dos proporciones y use este intervalo de confianza para determinar si existe una diferencia estadísticamente significativa en las proporciones de los números que no aparecen en el directorio entre las dos ciudades.

b) Encuentre un intervalo de confianza de 90% para la diferencia de las dos proporciones y use este intervalo de confianza para determinar si existe una diferencia estadísticamente significativa en las proporciones de los números que no aparecen en el directorio entre las dos ciudades.

c) Suponga que se duplican todos los números de la descripción del problema. Es decir, 774 residentes de 3 000 muestreados en Phoenix y 620 residentes de 2 400 de Scottsdale, tenían números telefónicos que no aparecían en el directorio. Repita los incisos a) y b) y comente el efecto sobre los resultados de aumentar el tamaño de la muestra sin cambiar las proporciones.

9-69. En una muestra aleatoria de 200 residentes de Phoenix que manejan un automóvil nacional, 165 dijeron usar regularmente el cinturón de seguridad, mientras que otra muestra de 250 residentes de Phoenix que manejan un automóvil extranjero reveló que 198 usan regularmente el cinturón de seguridad.

a) Aplique un procedimiento de prueba de hipótesis para determinar si hay una diferencia estadísticamente significativa en el uso del cinturón de seguridad entre los

conductores de automóviles nacionales y extranjeros. Fije la probabilidad de un error tipo I en 0.05.

b) Aplique un procedimiento de prueba de hipótesis para determinar si hay una diferencia estadísticamente significativa en el uso del cinturón de seguridad entre los conductores de automóviles nacionales y extranjeros. Fije la probabilidad de un error tipo I en 0.1.

c) Compare las respuestas de los incisos a) y b) y explique por qué son iguales o diferentes.

d) Suponga que se duplican todos los números de la descripción del problema. Es decir, en una muestra aleatoria de 400 residentes de Phoenix que manejan un automóvil nacional, 330 dijeron usar regularmente el cinturón de seguridad, mientras que otra muestra de 500 residentes de Phoenix que manejan un automóvil extranjero reveló que 396 usan regularmente el cinturón de seguridad. Repita los incisos a) y b) y comente el efecto sobre los resultados de aumentar el tamaño de la muestra sin cambiar las proporciones.

9-70. Considérese el ejercicio anterior, en el que se resumen los datos recogidos de conductores acerca del uso del cinturón de seguridad.

a) ¿Considera usted que hay alguna razón para no dar credibilidad a estos datos? Explique su respuesta.

b) ¿Es razonable usar los resultados de la prueba de hipótesis del problema anterior para hacer una inferencia acerca de la diferencia en la proporción del uso del cinturón de seguridad

i) de las esposas de estos conductores de automóviles nacionales y extranjeros? Explique su respuesta.

ii) de los hijos de estos conductores de automóviles nacionales y extranjeros? Explique su respuesta.

iii) de todos los conductores de automóviles nacionales y extranjeros? Explique su respuesta.

iv) de todos los conductores de camiones nacionales y extranjeros? Explique su respuesta.

9-71. Considérese el ejemplo 9-14 del texto.

a) Redefina los parámetros de interés para que sea la proporción de lentes que no son satisfactorios después del pulido rotatorio con los fluidos de pulido 1 o 2. Pruebe la hipótesis de que las dos soluciones de pulido producen resultados diferentes usando $\alpha = 0.01$.

b) Compare la respuesta del inciso a) con la del ejemplo. Explique por qué son iguales o diferentes.

9-72. Considérese el ejemplo 9-14 del texto, y recuerde que las hipótesis de interés son H_0: $p_1 = p_2$ contra H_1: $p_1 \neq p_2$. Quiere usarse $\alpha = 0.01$. Suponga que si $p_1 = 0.9$ y $p_2 = 0.6$ quiere detectarse este hecho con una alta probabilidad, digamos al menos 0.9. ¿Qué tamaños de las muestras se necesitan para alcanzar este objetivo?

9-73. Al fabricante de un nuevo analgésico en tabletas le gustaría demostrar que su producto trabaja el doble de rápido que el producto de la competencia. Específicamente, le gustaría probar

$$H_0 : \mu_1 = 2\mu_2$$
$$H_1 : \mu_1 > 2\mu_2$$

donde μ_1 es el tiempo promedio de absorción del producto de la competencia y μ_2 es el tiempo promedio de absorción del nuevo producto. Suponiendo que las varianzas σ_1^2 y σ_2^2 son conocidas, desarrolle un procedimiento para probar esta hipótesis.

9-74. Suponga que se está probando H_0: $\mu_1 = \mu_2$ contra H_1: $\mu_1 \neq \mu_2$, y que se planea usar el mismo tamaño para las muestras de las dos poblaciones. Se supone que ambas poblaciones son normales con varianzas desconocidas pero iguales. Si se usa $\alpha = 0.05$ y si la verdadera media $\mu_1 = \mu_2 + \sigma$, ¿qué tamaño de la muestra debe usarse para que la potencia de esta prueba sea de al menos 0.90?

9-75. Se realizó un estudio de economía de combustible para dos automóviles alemanes, Mercedes y Volskwagen. Se seleccionó un vehículo de cada marca y se observó el rendimiento en millas por galón para 10 tanques de combustible en cada auto. Los datos son los siguientes:

Mercedes		Volkswagen	
24.7	24.9	41.7	42.8
24.8	24.6	42.3	42.4
24.9	23.9	41.6	39.9
24.7	24.9	39.5	40.8
24.5	24.8	41.9	29.6

a) Construya una gráfica de probabilidad normal para cada serie de datos. Con base en estas gráficas, ¿es razonable suponer que cada serie se sacó de una población normal?

b) Suponga que se determinó que la observación más baja de los datos del Mercedes se registró incorrectamente y que debería ser 24.6. Además, la observación más baja de los datos del Volkswagen también estuvo equivocada y debería ser 39.6. Construya de nuevo las gráficas de probabilidad normal para cada serie de datos con los valores corregidos. Con base en estas nuevas gráficas, ¿es razonable suponer que cada serie se sacó de una población normal?

c) Compare las respuestas de los incisos a) y b) y comente el efecto de estas observaciones equivocadas sobre el supuesto de normalidad.

d) Usando los datos corregidos del inciso b) y un intervalo de confianza de 95%, ¿hay evidencia para soportar que la variabilidad en el rendimiento en millas por galón es más alto para un Volkswagen que para un Mercedes?

9-76. Considérese de nuevo el estudio de la economía del combustible del ejercicio complementario 9-75. Resuelva de nuevo el inciso d) de este problema usando un procedimiento de prueba de hipótesis adecuado. ¿Se llegó a la misma respuesta obtenida originalmente? ¿Por qué?

9-77. Considérense los agentes de expansión de espuma para combatir incendios que se investigaron en el ejercicio 9-16, donde se registraron cinco observaciones de cada agente. Suponga que si el agente 1 produce una expansión media que difiere de la expansión media del agente 2 por 1.5, entonces querría rechazarse la hipótesis nula con una probabilidad de al menos 0.95.

a) ¿Qué tamaño de la muestra se necesita?

b) ¿Considera usted que el tamaño de la muestra original del ejercicio 9-16 fue apropiado para detectar esta diferencia? Explique su respuesta.

9-78. Se realizó un experimento para comparar la capacidad de llenado del equipo de embotellado en dos fábricas vinícolas diferentes. Se seleccionaron al azar y se midieron diez botellas de Pinot Noir de Ridgecrest Vineyards, junto con 10 botellas de Pinot Noir de Valley View Vineyards. Los datos son los siguientes (el volumen de llenado está en ml):

Ridgecrest				Valley View			
755	751	752	753	756	754	757	756
753	753	753	754	755	756	756	755
752	751			755	756		

a) ¿Qué supuestos se necesitan para realizar un procedimiento de prueba de la igualdad de la medias de estos datos? Verifique estos supuestos.

b) Realice el procedimiento de prueba de hipótesis adecuado para determinar si los datos apoyan la afirmación de que ambas fábricas vinícolas llenan las botellas con el mismo volumen promedio.

9-79. Considérese el ejercicio complementario 9-78. Suponga que la diferencia real del volumen de llenado promedio es a lo más de 2 onzas líquidas; ¿los tamaños de las muestras de 10 botellas de cada fábrica vinícola proporcionan una buena capacidad de detección cuando $\alpha = 0.05$? Explique su respuesta.

9-80. Una máquina Rockwell para probar la dureza presiona una punta en una muestra de acero

para prueba, y la profundidad de la depresión resultante se usa para indicar la dureza. Se comparan dos puntas diferentes para determinar si producen las mismas lecturas de dureza en la escala C de Rockwell. Se usan nueve muestras de acero, probándose ambas puntas en cada una de ellas. Los datos se muestran en la tabla siguiente:

Muestra de acero	Punta 1	Punta 2
1	47	46
2	42	40
3	43	45
4	40	41
5	42	43
6	41	41
7	45	46
8	45	46
9	49	48

a) Enuncie cualquier supuesto necesario para probar la afirmación de que ambas puntas producen las mismas lecturas de dureza en la escala C de Rockwell. Verifique los supuestos con los cuales estableció usted los datos.

b) Aplique un método estadístico adecuado para determinar si los datos apoyan la afirmación de que la diferencia de las lecturas de dureza en la escala C de Rockwell de las dos puntas difiere significativamente de cero.

c) Suponga que si las dos puntas difieren en las lecturas de la dureza media a lo más por 1.0, se quiere que la potencia de la prueba sea de al menos 0.9. Para $\alpha = 0.01$, ¿cuántas muestras de acero deberán usarse en la prueba?

9-81. Pueden usarse dos medidores para registrar la profundidad del material de baño en una celda Hall usada en la fundición de aluminio. El mismo operador usa una vez cada medidor en 15 celdas.

Celda	Medidor 1	Medidor 2
1	46 pulg	47 pulg
2	50	53
3	47	45
4	53	50
5	49	51
6	48	48
7	53	54
8	56	53
9	52	51
10	47	45
11	49	51
12	45	45
13	47	49
14	46	43
15	50	51

a) Enuncie cualquier supuesto necesario para probar la afirmación de que ambos medidores producen la misma lectura media de la profundidad del baño. Verifique los supuestos con los cuales estableció usted los datos.

b) Aplique un procedimiento estadístico adecuado para determinar si los datos apoyan la afirmación de que los dos medidores producen diferentes lecturas medias de la profundidad del baño.

c) Suponga que si los dos medidores difieren en las lecturas medias de la profundidad del baño a lo más por 1.65 pulgadas, se quiere que la potencia de la prueba sea de al menos 0.8. Para $\alpha = 0.01$, ¿cuántas celdas deberán usarse?

EJERCICIOS PARA DESARROLLAR EL INTELECTO

9-82. Pueden usarse tres pesticidas diferentes para controlar las plagas de las vides. Se piensa que el pesticida 3 es más efectivo que los otros dos. En un viñedo particular, se seleccionan tres plantíos diferentes de uvas Pinot Noir para estudio. Se obtienen los siguientes resultados del rendimiento:

Pesticida	$\bar{x}_i$ (Bushels/ planta)	s_i	n_i (Número de plantas)
1	4.6	0.7	100
2	5.2	0.6	120
3	6.1	0.8	130

Si μ_i es el verdadero rendimiento medio después del tratamiento con el pesticida i, entonces el interés se encuentra en la cantidad

$$\mu = \frac{1}{2}(\mu_1 + \mu_2) - \mu_3$$

que mide la diferencia de los rendimientos medios entre los pesticidas 1 y 2 y el pesticida 3. Si los tamaños de las muestras n_i son grandes, el estimador (por ejemplo, $\hat{\mu}$) que se obtiene al sustituir cada μ_i individual por $\bar{X}_i$ es aproximadamente normal.

a) Encuentre un intervalo de confianza aproximado del $100(1-\alpha)$ por ciento de muestras grandes para μ.

b) ¿Estos datos apoyan la afirmación de que el pesticida 3 es más efectivo que los otros dos? Use $\alpha = 0.05$ para determinar la respuesta.

9-83. Suponga que quiere probarse H_0: $\mu_1 = \mu_2$ contra H_1: $\mu_1 \neq \mu_2$, donde σ_1^2 y σ_2^2 son conocidas. Debe terminarse el tamaño total de la muestra N, y la asignación del número de observaciones de las dos poblaciones, tales que $n_1 + n_2 = N$ se hará con base en el costo. Si el costo del muestreo de las poblaciones 1 y 2 son C_1 y C_2, respectivamente, encuentre los tamaños de las muestras con el costo mínimo que produzcan una varianza especificada para la diferencia de las medias muestrales.

9-84. Suponga que quiere probarse la hipótesis H_0: $\mu_1 = \mu_2$ contra H_1: $\mu_1 \neq \mu_2$, donde σ_1^2 y σ_2^2 son conocidas. Puede tomarse un total de $n_1 + n_2 = N$ observaciones. ¿Cómo deben asignarse estas observaciones en las dos poblaciones para maximizar la probabilidad de que H_0 se rechace si H_1 es verdadera y $\mu_1 - \mu_2 = \Delta \neq 0$?

9-85. Suponga que quiere probarse H_0: $\mu = \mu_0$ contra H_1: $\mu \neq \mu_0$, donde la población es normal con σ conocida. Sea $0 < \varepsilon < \alpha$, y la región crítica se define de tal modo que H_0 se rechaza si $z_0 > z_\varepsilon$ o si $z_0 < -z_{\alpha-\varepsilon}$, donde z_0 es el valor del estadístico de prueba usual para estas hipótesis.

a) Demuestre que la probabilidad del error tipo I para esta prueba es α.

b) Suponga que la verdadera media es $\mu_1 = \mu_0 + \Delta$. Deduzca una expresión para β en la prueba anterior.

9-86. Construya una serie de datos para los que el estadístico de la prueba t pareada sea muy grande, lo cual indicaría que cuando se aplique este análisis, las dos medias poblacionales son diferentes, pero t_0 de la prueba t de dos muestras es muy pequeño, de tal modo que el análisis incorrecto indicaría que no hay una diferencia significativa entre las medias.

9-87. Deduzca una expresión para β para la prueba de la varianza de una distribución normal. Suponer que se ha especificado la hipótesis alternativa de dos colas.

9-88. Deduzca una expresión para β en la prueba de la igualdad de las varianzas de dos distribuciones normales. Suponga que se especifica la hipótesis alternativa de dos colas.

10 Regresión lineal simple y correlación

PLAN GENERAL DEL CAPÍTULO

10-1 MODELOS EMPÍRICOS

10-2 REGRESIÓN LINEAL SIMPLE

10-3 PROPIEDADES DE LOS ESTIMADORES DE MÍNIMOS CUADRADOS Y LA ESTIMACIÓN DE σ^2

10-4 ABUSOS COMUNES DE LA REGRESIÓN

10-5 PRUEBAS DE HIPÓTESIS EN LA REGRESIÓN LINEAL SIMPLE

- 10-5.1 Uso de las pruebas t
- 10-5.2 Enfoque del análisis de varianza para probar la significación de una regresión

10-6 INTERVALOS DE CONFIANZA

- 10-6.1 Intervalos de confianza para la pendiente y la ordenada al origen
- 10-6.2 Intervalo de confianza para la respuesta media

10-7 PREDICCIÓN DE NUEVAS OBSERVACIONES

10-8 EVALUANDO LA ADECUACIÓN DEL MODELO DE REGRESIÓN

- 10-8.1 Análisis de residuales
- 10-8.2 Coeficiente de determinación (R^2)
- 10-8.3 Prueba de falta de ajuste

10-9 TRANSFORMACIONES A UNA LÍNEA RECTA

10-10 CORRELACIÓN

10-1 MODELOS EMPÍRICOS

En muchos problemas, dos o más variables tienen una relación inherente, y es necesario explorar la naturaleza de la misma. El **análisis de regresión** es una técnica estadística para modelar e investigar la relación entre dos o más variables. Por ejemplo, suponga que el rendimiento del producto en un proceso químico está relacionado con la temperatura de operación del proceso. Puede usarse un análisis de regresión para construir un modelo que permita predecir el rendimiento en un nivel de temperatura dado. Este modelo también puede usarse para optimizar el proceso, como por ejemplo al encontrar el nivel de temperatura que maximice el rendimiento, o para fines de control del proceso.

Para ilustrar el concepto, considérense los datos de la tabla 10-1. En esta tabla, y es la pureza del oxígeno producido en un proceso de destilación química, y x es el porcentaje de hidrocarburos que están presentes en el condensador principal de la unidad de destilación. En la figura 10-1 se presenta un **diagrama de dispersión** de los datos de la tabla 10-1. Se trata de una gráfica en la que cada par (x_i, y_i) está representado por un punto graficado en un sistema de coordenadas bidimensionales. Al inspeccionar este diagrama de dispersión se observa que, aun cuando ninguna curva simple pasará exactamente por todos los puntos, hay claros indicios de que los puntos se encuentran dispersos aleatoriamente alrededor de una línea recta. Por lo tanto,

Tabla 10-1 Niveles de oxígeno de hidrocarburos

Número de observación	Nivel de hidrocarburos x (%)	Pureza y (%)
1	0.99	90.01
2	1.02	89.05
3	1.15	91.43
4	1.29	93.74
5	1.46	96.73
6	1.36	94.45
7	0.87	87.59
8	1.23	91.77
9	1.55	99.42
10	1.40	93.65
11	1.19	93.54
12	1.15	92.52
13	0.98	90.56
14	1.01	89.54
15	1.11	89.85
16	1.20	90.39
17	1.26	93.25
18	1.32	93.41
19	1.43	94.98
20	0.95	87.33

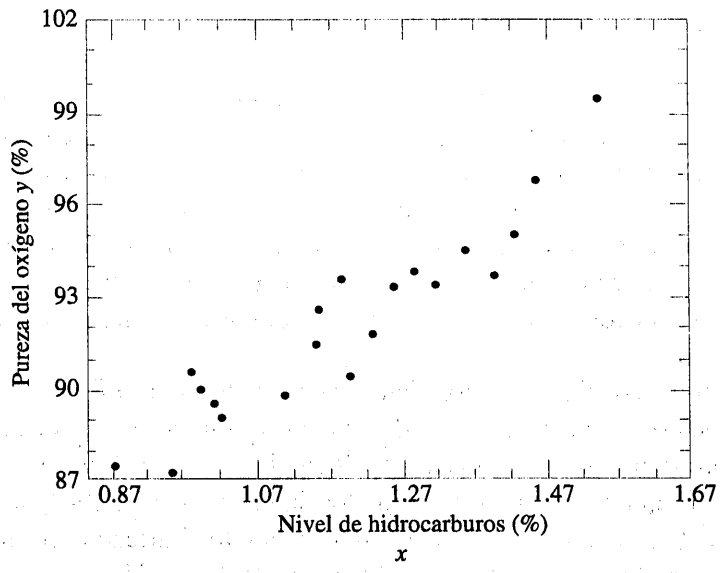


Figura 10-1 Diagrama de dispersión de la pureza del oxígeno contra el nivel de hidrocarburos de la tabla 10-1.

quizás sea razonable suponer que la media de la variable aleatoria Y se relaciona con x por la siguiente relación lineal:

$$E(Y|x) = \mu_{Y|x} = \beta_0 + \beta_1 x$$

donde a la pendiente y a la ordenada al origen se les llama los **coeficientes de regresión**. Aun cuando la media de Y es una función lineal de x, el valor real y observado no está exactamente sobre una línea recta. La forma apropiada de generalizar lo anterior para llegar a un **modelo lineal probabilístico** es suponer que el valor esperado de Y es una función lineal de x, pero que el valor real de Y para un valor fijo de x se determina por la función del valor medio (el modelo lineal) más un término del error aleatorio, por ejemplo,

$$Y = \beta_0 + \beta_1 x + \varepsilon \tag{10-1}$$

donde ε es el término del error aleatorio. A este modelo se le llamará **modelo de regresión lineal simple**, porque sólo tiene una variable independiente o **regresor** (**variable de regresión**). En ocasiones, un modelo como éste surgirá de una relación teórica. En otras ocasiones, no se contará con un conocimiento teórico de la relación entre x y y, y la elección del modelo se basará en la inspección de un diagrama de dispersión, como se hizo con los datos de la pureza del oxígeno. Por lo tanto, el modelo de regresión se considera un **modelo empírico**.

Para profundizar más en este modelo, suponer que se puede fijar el valor de x y observar el valor de la variable aleatoria Y. Entonces, si x está fija, el componente aleatorio ε del miembro derecho del modelo de la ecuación 10-1 determina las propiedades de Y. Suponga que la media y la varianza de ε son 0 y σ^2, respectivamente. Entonces

$$E(Y|x) = E(\beta_0 + \beta_1 x + \varepsilon) = \beta_0 + \beta_1 x + E(\varepsilon) = \beta_0 + \beta_1 x$$

Obsérvese que esta es la misma relación que se escribió empíricamente en un principio a partir de la inspección del diagrama de dispersión de la figura 10-1. La varianza de Y dada x es

$$V(Y|x) = V(\beta_0 + \beta_1 x + \varepsilon) = V(\beta_0 + \beta_1 x) + V(\varepsilon) = 0 + \sigma^2 = \sigma^2$$

Por tanto, el verdadero modelo de regresión $\mu_{Y|x} = \beta_0 + \beta_1 x$ es una recta de valores medios; es decir, la altura de la recta de regresión en cualquier valor de x es el valor esperado de Y para esa x. La pendiente, β_1, puede interpretarse como el cambio en la media de Y para un cambio unitario en x. Además, la variabilidad de Y en un valor particular de x está determinada por la varianza del error, σ^2. Esto implica que hay una distribución de valores Y en cada x y que la varianza de esta distribución es la misma en cada x.

Por ejemplo, suponga que el verdadero modelo de regresión que relaciona la pureza del oxígeno con el nivel de hidrocarburos es $\mu_{Y|x} = 75 + 15x$, y suponga que la varianza es $\sigma^2 = 2$. En la figura 10-2 se ilustra esta situación. Obsérvese que se ha usado una distribución normal para describir la variación aleatoria de ε. Puesto que Y es la suma de una constante $\beta_0 + \beta_1 x$ (la media) y de una variable aleatoria que tiene una distribución normal, Y es una variable aleatoria con una distribución normal. La varianza σ^2 determina la variabilidad en las observaciones Y de la pureza del oxígeno. Por tanto, cuando σ^2 es pequeña, los valores observados de Y estarán cerca de la recta, y cuando σ^2 es grande, los valores observados de Y pueden apartarse considerablemente de la recta. Puesto que σ^2 es constante, la variabilidad de Y en cualquier valor de x es la misma.

El modelo de regresión describe la relación entre la pureza del oxígeno Y y el nivel de hidrocarburos x. Por tanto, para cualquier valor del nivel de hidrocarburos, la pureza del oxígeno tiene una distribución normal con media $75 + 15x$ y varianza 2. Por ejemplo, si $x = 1.25$, entonces Y tiene el valor medio $\mu_{Y|x} = 75 + 15(1.25) = 93.75$ y varianza 2.

En la mayoría de los problemas del mundo real, no se conocerán los valores de la ordenada al origen y la pendiente (β_0, β_1) ni el de la varianza del error σ^2, por lo que deberán estimarse a partir de los datos muestrales. El **análisis de regresión** es una colección de herramientas estadísticas para encontrar las estimaciones de los parámetros del modelo de regresión. Posteriormente, esta ecuación de regresión o modelo de regresión ajustado se usa por lo general para predecir observaciones futuras de Y, o para estimar la respuesta media en un nivel particular de x. Como ilustración, un ingeniero químico podría interesarse en estimar la pureza media del

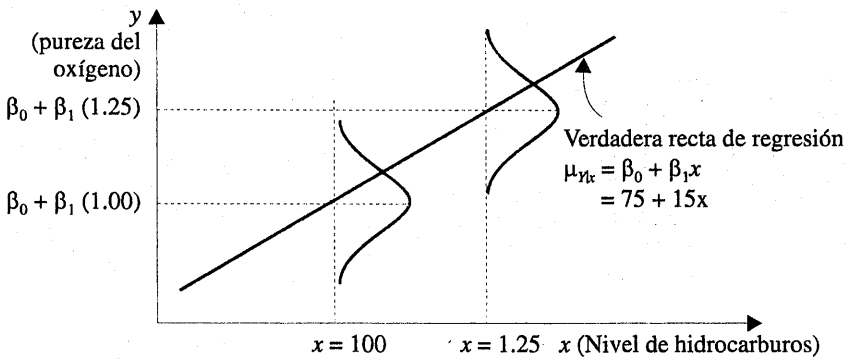


Figura 10-2 La distribución de Y para un valor dado de x para los datos, pureza del oxígeno-hidrocarburos.

oxígeno producido cuando el nivel de hidrocarburos es $x = 1.25\%$. En este capítulo se consideran estos procedimientos y las aplicaciones del modelo de regresión lineal simple. En el capítulo 11 se analizan los modelos de regresión lineal múltiple en los que interviene más de un regresor (variable de regresión).

10-2 REGRESIÓN LINEAL SIMPLE

En el caso de la **regresión lineal simple** se considera un *solo regresor (variable de regresión)* o *predictor x*, y una variable dependiente o *variable de respuesta Y*. Suponga que la verdadera relación entre Y y x es una recta y que la observación Y en cada nivel de x es una variable aleatoria. Como ya se señaló, el valor esperado de Y para cada valor de x es

$$E(Y|x) = \beta_0 + \beta_1 x$$

donde la ordenada al origen β_0 y la pendiente β_1 son coeficientes de regresión desconocidos. Se supone que cada observación, Y, puede describirse con el modelo

$$Y = \beta_0 + \beta_1 x + \varepsilon \tag{10-2}$$

donde ε es un error aleatorio con media cero y varianza σ^2. Se supone asimismo que los errores aleatorios correspondientes a diferentes observaciones son variables aleatorias no correlacionadas.

Suponga que se tienen n pares de observaciones $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. En la figura 10-3 se presenta una gráfica de dispersión típica de los datos observados y un candidato para la recta de regresión estimada. Las estimaciones de β_0 y β_1 deberán dar como resultado una recta que sea (en algún sentido) el "mejor ajuste" para los datos. El científico alemán Karl Gauss (1777-1855) propuso estimar los parámetros β_0 y β_1 de la ecuación 10-2 a fin de minimizar la suma de los cuadrados de las desviaciones verticales de la figura 10-3.

A este criterio para estimar los coeficientes de regresión se le llama el **método de mínimos cuadrados**. Usando la ecuación 10-2, las n observaciones de la muestra pueden expresarse como

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad i = 1, 2, \ldots, n \tag{10-3}$$

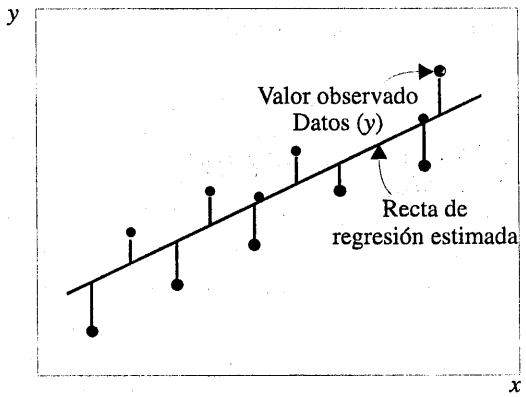


Figura 10-3 Desviaciones de los datos del modelo de regresión estimado.

y la suma de los cuadrados de las desviaciones de las observaciones de la verdadera recta de regresión es

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 + \beta_1 x_i)^2 \tag{10-4}$$

Los estimadores de mínimos cuadrados de β_0 y β_1, por ejemplo $\hat{\beta}_0$ y $\hat{\beta}_1$, deben satisfacer

$$\begin{aligned} \frac{\partial L}{\partial \beta_0}\Big|_{\hat{\beta}_0, \hat{\beta}_1} &= -2\sum_{i=1}^{n} (y_i - \hat{\beta}_0 + \hat{\beta}_1 x_i) = 0 \\ \frac{\partial L}{\partial \beta_1}\Big|_{\hat{\beta}_0, \hat{\beta}_1} &= -2\sum_{i=1}^{n} (y_i - \hat{\beta}_0 + \hat{\beta}_1 x_i) x_i = 0 \end{aligned} \tag{10-5}$$

Al simplificar estas dos ecuaciones se obtiene

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} (y_i) \\ \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} y_i x_i \end{aligned} \tag{10-6}$$

A las ecuaciones 10-6 se les llama las **ecuaciones normales de mínimos cuadrados**. Al resolver las ecuaciones normales se obtienen los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$.

Definición

Las **estimaciones de mínimos cuadrados** de la ordenada al origen y la pendiente del modelo de regresión lineal simple son

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{10-7}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \dfrac{\left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} x_i\right)}{n}}{\sum_{i=1}^{n} x_i^2 - \dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}} \tag{10-8}$$

donde $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$ y $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$.

Por lo tanto, la **recta de regresión ajustada** o **estimada** es

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x \tag{10-9}$$

Obsérvese que cada par de observaciones satisface la relación

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i, \qquad i = 1, 2, \ldots, n$$

donde a $e_i = y_i - \hat{y}_i$ se le llama el **residual**. El residual describe el error del ajuste del modelo en la observación i, y_i. Más adelante en este capítulo se usarán los residuales para proporcionar información acerca de la adecuación del modelo ajustado.

En lo que a la notación se refiere, en ocasiones es conveniente asignar símbolos especiales al numerador y al denominador de la ecuación 10-8. Dados los datos $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, sean

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} \tag{10-10}$$

y

$$S_{xy} = \sum_{i=1}^{n} y_i (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n} \tag{10-11}$$

EJEMPLO 10-1

Se ajustará un modelo de regresión lineal simple a los datos de la pureza del oxígeno de la tabla 10-1. Pueden calcularse las siguientes cantidades:

$$n = 20 \quad \sum_{i=1}^{20} x_i = 23.92 \quad \sum_{i=1}^{20} y_i = 1\,843.21 \quad \bar{x} = 1.20 \quad \bar{y} = 92.16$$

$$\sum_{i=1}^{20} y_i^2 = 170\,044.53 \quad \sum_{i=1}^{20} x_i^2 = 29.29 \quad \sum_{i=1}^{20} x_i y_i = 2\,214.66$$

$$S_{xx} = \sum_{i=1}^{20} x_i^2 - \frac{\left(\sum_{i=1}^{20} x_i\right)^2}{20} = 29.29 - \frac{(23.92)^2}{20} = 0.68$$

y

$$S_{xy} = \sum_{i=1}^{20} x_i y_i - \frac{\left(\sum_{i=1}^{20} x_i\right)\left(\sum_{i=1}^{20} y_i\right)}{20} = 2\,214.66 - \frac{(23.92)(1\,843.21)}{20} = 10.18$$

Por lo tanto, las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen son

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{10.18}{0.68} = 14.97$$

y

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 92.16 - (14.97)1.20 = 74.20$$

El modelo de regresión lineal simple ajustado es

$$\hat{y} = 74.20 + 14.97x$$

Este modelo se grafica en la figura 10-4, junto con los datos muestrales. Se usaron dos cifras decimales en el cálculo de estos coeficientes de regresión.

Los programas de software de computadora son de uso generalizado para modelar regresiones. Estos programas por lo general manejan más cifras decimales en los cálculos. En la tabla 10-2 se muestra una parte de la salida del programa Minitab para este problema. Obsérvese que las estimaciones de la pendiente y la ordenada al origen se reportan con tres cifras decimales, y no coinciden exactamente con los cálculos manuales porque en éstos se usaron menos cifras decimales en las operaciones aritméticas que con la computadora. Sería necesario manejar al menos cuatro cifras decimales en los cálculos manuales para que coincidan con la salida de computadora. En secciones posteriores se explicará el resto de la información proporcionada en esta salida de computadora.

Al utilizar el modelo de regresión del ejemplo 10-1, se puede predecir una pureza del oxígeno de $\hat{y} = 89.17\%$ cuando el nivel de hidrocarburos es $x = 1.00\%$. La pureza 89.17% puede

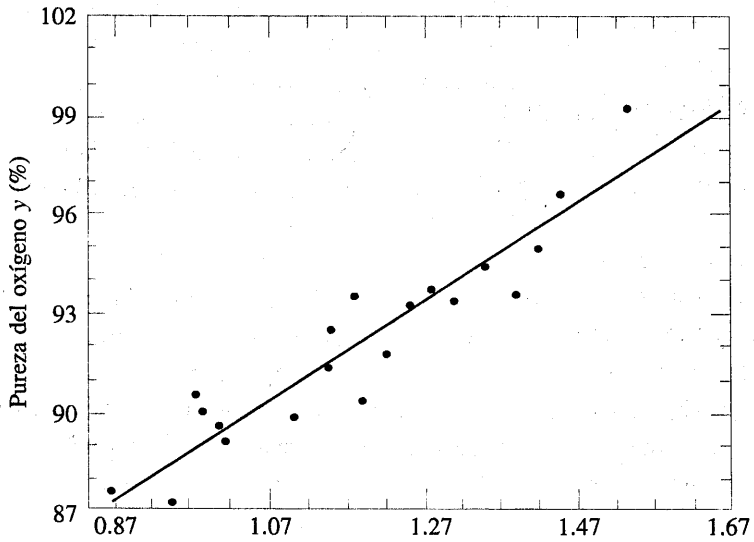


Figura 10-4 Gráfica de dispersión de la pureza del oxígeno y contra el nivel de hidrocarburos x y el modelo de regresión $\hat{y} = 74.20 + 14.97x$.

Tabla 10-2 Salida del programa Minitab para los datos de la pureza del oxígeno del ejemplo 10-1

Análisis de regresión

La ecuación de regresión es
$y = 74.3 + 14.9x$

Predictor	Coef	Desv. Est.	T	P
Constante	74.283	1.593	46.62	0.000
x	14.947	1.317	11.35	0.000

S = 1.087 R-Sq = 87.7% R-Sq(ajustada) = 87.1%

Análisis de varianza

Fuente	DF	SS	MS	F	P
Regresión	1	152.13	152.13	128.86	0.000
Error	18	21.25	1.18		
Total	19	173.38			

interpretarse como una estimación de la verdadera pureza media de la población cuando x = 1.00%, o como una estimación de una nueva observación cuando x = 1.00%. Desde luego, estas estimaciones están sujetas a error; es decir, es improbable que una observación futura de la pureza sea exactamente 89.17% cuando el nivel de hidrocarburos es 1.00%. En secciones posteriores se verá cómo usar intervalos de confianza y predicción de intervalos para describir el error de una estimación hecha a partir de un modelo de regresión.

EJERCICIOS DE LA SECCIÓN 10-2

10-1. En un artículo de *Concrete Research* ("Características del concreto cerca de la superficie: Permeabilidad intrínseca" (Vol. 41), se presentaron los datos sobre la resistencia a la compresión x y la permeabilidad intrínseca y de varias mezclas y curados de concreto. Las cantidades resumidas son $n = 14$, $\Sigma y_i = 572$, $\Sigma y_i^2 = 23\ 530$, $\Sigma x_i = 43$, $\Sigma x_i^2 = 157.42$ y $\Sigma x_i y_i = 1\ 697.80$. Suponga que las dos variables se relacionan de acuerdo con el modelo de regresión lineal simple.

a) Calcule las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen.

b) Use la ecuación de la recta ajustada para predecir la permeabilidad que se observaría cuando la resistencia a la compresión es $x = 4.3$.

c) Dé una estimación puntual de la permeabilidad media cuando la resistencia a la compresión es $x = 3.7$.

d) Suponga que el valor observado de la permeabilidad para $x = 3.7$ es $y = 46.1$. Calcule el valor del residuo correspondiente.

10-2. Se utilizaron métodos de regresión para analizar los datos de un estudio para investigar la relación entre la temperatura superficial de una carretera (x) y la deflexión del pavi-

mento (y). Las cantidades resumidas fueron $n = 20$, $\Sigma y_i = 12.75$, $\Sigma y_i^2 = 8.86$, $\Sigma x_i = 1\,478$, $\Sigma x_i^2 = 143\,215.8$ y $\Sigma x_i y_i = 1\,083.67$.

a) Calcule las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen. Grafique la recta de regresión.

b) Use la ecuación de la recta ajustada para predecir la deflexión del pavimento que se observaría cuando la temperatura superficial es 85°F.

c) ¿Cuál es la deflexión media del pavimento cuando la temperatura superficial es 90°F?

d) ¿Qué cambio en la deflexión media del pavimento se esperaría para un cambio de 1°F en la temperatura superficial?

10-3. Considérese el modelo de regresión desarrollado en el ejercicio 10-2.

a) Suponga que la temperatura se mide en °C en vez de en °F. Escriba el nuevo modelo de regresión que resulta.

b) ¿Qué cambio en la deflexión esperada del pavimento se asocia con un cambio de 1°C en la temperatura superficial?

10-4. Montgomery y Peck presentan los datos del desempeño de los 28 equipos de la Liga Nacional de Futbol Americano (NFL, por sus siglas en inglés) en 1976. Se piensa que el número de juegos ganados (y) se relaciona con el número de yardas ganadas en el campo por los oponentes (x). Los datos se muestran en la tabla siguiente.

Equipos	Juegos ganados (y)	Yardas corridas en el campo de los oponentes (x)
Washington	10	2 205
Minnesota	11	2 096
Nueva Inglaterra	11	1 847
Oakland	13	1 903
Pittsburg	10	1 457
Baltimore	11	1 848
Los Ángeles	10	1 564
Dallas	11	1 821
Atlanta	4	2 577
Búffalo	2	2 476
Chicago	7	1 984
Cincinnati	10	1 917
Cleveland	9	1 761
Denver	9	1 709
Detroit	6	1 901
Green Bay	5	2 288
Houston	5	2 072
Kansas City	5	2 861
Miami	6	2 411
Nueva Orleans	4	2 289
Gigantes de Nueva York	3	2 203
Jets de Nueva York	3	2 592
Filadelfia	4	2 053
San Luis	10	1 979
San Diego	6	2 048
San Francisco	8	1 786
Seattle	2	2 876
Tampa Bay	0	2 560

a) Calcule las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen. Grafique el modelo de regresión.

b) Encuentre una estimación del número promedio de juegos ganados si los oponentes pueden limitarse a 1 800 yardas adentradas en el campo.

c) ¿Qué cambio en el número esperado de juegos ganados se asocia con una disminución de 100 yardas corridas en el campo de un oponente?

d) Para incrementar en uno el número promedio de juegos ganados, ¿qué disminución de las yardas corridas en el campo debe generar la defensiva?

e) Dado que $x = 1\,917$ yardas (Cincinnati), encuentre el valor ajustado de y y el residuo correspondiente.

10-5. Un artículo de *Technometrics* de S.C. Narula y J.F. Wellington ("Predicción, regresión lineal, y suma mínima de errores relativos", Vol. 19) presenta los datos sobre el precio de venta y los impuestos anuales de 24 casas. Los datos se muestran en la tabla siguiente:

Precio de venta/1 000	Impuestos (locales, educación, municipales)/1 000
25.9	4.9176
29.5	5.0208
27.9	4.5429
25.9	4.5573
29.9	5.0597
29.9	3.8910
30.9	5.8980
28.9	5.6039
35.9	5.8282
31.5	5.3003
31.0	6.2712
30.9	5.9592
30.0	5.0500
36.9	8.2464
41.9	6.6969
40.5	7.7841
43.9	9.0384
37.5	5.9894
37.9	7.5422
44.5	8.7951
37.9	6.0831
38.9	8.3607
36.9	8.1400
45.8	9.1416

a) Suponiendo que un modelo de regresión lineal simple es apropiado, obtenga el ajuste de mínimos cuadrados que relacione el precio de venta con los impuestos pagados.

b) Encuentre el precio de venta promedio, dado que los impuestos pagados son $x = 7.50$.

c) Calcule el valor ajustado de y que corresponde a $x = 5.8980$. Encuentre el residual correspondiente.

d) Calcule el valor ajustado $\hat{y}_i$ de cada valor de x_i usado para ajustar el modelo. Después, construya una gráfica de $\hat{y}_i$ contra el valor observado correspondiente y_i y comente cómo debería verse esta gráfica si la relación entre y y x fuera una línea recta determinista (sin error aleatorio). ¿La gráfica que se obtuvo realmente indica que los impuestos pagados es una variable regresora efectiva para predecir el precio de venta?

10-6. Se piensa que el número de libras de vapor consumidas mensualmente por una planta química se relaciona con la temperatura ambiente promedio (en °F) de ese mes. En la tabla siguiente se muestran la temperatura y el consumo anual:

Mes	Temperatura	Consumo/1 000
Ene.	21	185.79
Feb.	24	214.47
Mar.	32	288.03
Abr.	47	424.84
May.	50	454.58
Jun.	59	539.03
Jul.	68	621.55
Ago.	74	675.06
Sep.	62	562.03
Oct.	50	452.93
Nov.	41	369.95
Dic.	30	273.98

a) Suponiendo que un modelo de regresión lineal simple es apropiado, ajuste el modelo de regresión que relacione el consumo de vapor (y) con la temperatura promedio (x).

b) ¿Cuál es la estimación del consumo esperado de vapor cuando la temperatura promedio es 55°F?

c) ¿Qué cambio se espera en el consumo de vapor promedio cuando la temperatura mensual promedio cambia 1°F?

d) Suponga que la temperatura mensual promedio es 47°F. Calcule el valor ajustado de y y el residual correspondiente.

10-7. Los datos que se presentan en la tabla siguiente corresponden al rendimiento de gasolina

en carretera y la cilindrada del motor de una muestra de 20 automóviles.

Marca	Modelo	Millas por galón (en carretera)	Cilindraje del motor (pulgadas^3)
Acura	Legend	30	97
BMW	735i	19	209
Buick	Regal	29	173
Chevrolet	Cavalier	32	121
Chevrolet	Celebrity	30	151
Chrysler	Conquest	24	156
Dodge	Aries	30	135
Dodge	Dynasty	28	181
Ford	Escort	31	114
Ford	Mustang	25	302
Ford	Taurus	27	153
Ford	Tempo	33	90
Honda	Accord	30	119
Mazda	RX-7	23	80
Mercedes	260E	24	159
Mercury	Tracer	29	97
Nissan	Maxima	26	181
Oldsmobile	Cutlass	29	173
Plymouth	Laser	37	122
Pontiac	Grand Prix	29	173

a) Ajuste un modelo lineal simple que relacione las millas por galón en carretera (y) con el desplazamiento del motor (x) utilizando mínimos cuadrados.

b) Encuentre una estimación del rendimiento en millas promedio en carretera para un automóvil con un motor de 150 pulg^3 de cilindrada.

c) Obtenga el valor ajustado de y y el residual correspondiente para un automóvil, el Ford Escort, con un motor de 114 pulg^3 de cilindrada.

10-8. En un artículo de *Tappi Journal* se presentaron los datos de la concentración del licor verde Na_2S y la producción de una máquina papelera. Los datos (tomados de una gráfica) se muestran en la tabla siguiente:

Número de observación	Concentración del licor verde Na_2S (g/l)	Producción (toneladas/día)
1	40	825
2	42	830
3	49	890
4	46	895
5	44	890
6	48	910
7	46	915
8	43	960
9	53	990
10	52	1010
11	54	1012
12	57	1030
13	58	1050

a) Ajuste un modelo de regresión lineal simple con la concentración del licor verde Na_2S y y la producción x. Trace un diagrama de dispersión de los datos y el modelo ajustado de mínimos cuadrados resultante.

b) Encuentre el valor ajustado de y correspondiente a $x = 910$ y el residual asociado.

c) Encuentre la concentración media del licor verde Na_2S cuando la rapidez de producción es de 950 toneladas por día.

10-9. En un artículo de *Journal of Sound and Vibration* (Vol. 151, pp. 383-394) se describe un estudio para investigar la relación entre la exposición al ruido y la hipertensión. Los datos siguientes son representativos de los que se reportan en el artículo.

a) Trace un diagrama de dispersión de y (aumento de la presión sanguínea) contra x (nivel de presión del sonido). ¿Parece razonable un modelo de regresión lineal simple en esta situación?

b) Ajuste el modelo de regresión lineal simple utilizando mínimos cuadrados.

Aumento de la presión sanguínea provocado (mmHg)	Nivel de presión del sonido (dB)
1	60
0	63
1	65
2	70
5	70
1	70
4	80
6	80
2	80
3	80
5	85
4	89
6	90
8	90
4	90
5	90
7	94
9	100
7	100
6	100

c) Encuentre el aumento promedio predicho del nivel de presión sanguínea asociado con un nivel de presión del sonido de 85 dB.

10-10. En un artículo de *Wear* (Vol. 152, pp. 171-181) se presentan los datos del desgaste por rozamiento del acero dulce y la viscosidad del aceite. Los datos representativos, con x = viscosidad del aceite y y = volumen del desgaste (10^{-4} mm), son:

y	240	181	193	155	172
x	1.6	9.4	15.5	20.0	22.0
y	110	113	75	94	
x	35.5	43.0	40.5	33.0	

a) Construya un diagrama de dispersión de los datos. ¿Parecería plausible un modelo de regresión lineal simple?

b) Ajuste el modelo de regresión lineal simple usando mínimos cuadrados.

c) Prediga el desgaste por rozamiento cuando la viscosidad es $x = 30$.

d) Obtenga el valor ajustado de y cuando $x = 22.0$ y calcule el residual correspondiente.

10-11. En un artículo de *Journal of Environmental Engineering* (Vol. 115, pp. 608-619) se reportan los resultados de un estudio sobre la presencia de sodio y cloruros en corrientes superficiales de la parte central de Rhode Island. Los datos que se presentan a continuación corresponden a la concentración de cloruros y (en mg/l) y al área de carretera de la vertiente (en %).

y	4.4	6.6	9.7	10.6	10.8
x	0.19	0.15	0.57	0.70	0.67
y	10.9	11.8	12.1	14.3	14.7
x	0.63	0.47	0.70	0.60	0.78
y	15.0	17.3	19.2	23.1	27.4
x	0.81	0.78	0.69	1.30	1.05
y	27.7	31.8	39.5		
x	1.06	1.74	1.62		

a) Trace un diagrama de dispersión de los datos. ¿Parecería apropiado un modelo de regresión lineal simple en este caso?

b) Ajuste el modelo de regresión lineal simple usando el método de mínimos cuadrados.

c) Estime la concentración de cloruros media de una vertiente que tiene 1% del área de carretera.

d) Encuentre el valor ajustado que corresponde a $x = 0.47$ y el residual asociado.

10-12. Un motor de turbina se fabrica ensamblando dos tipos de cargas propulsoras, un mecanismo de ignición y un soporte. Se piensa que la resistencia al corte de la unidad ensamblada y es una función lineal de la antigüedad de la carga propulsora x cuan-

do se moldea el motor. En la tabla siguiente se muestran 20 observaciones.

Número de observación	Resistencia al corte y (psi)	Antigüedad x (semanas)
1	2 158.70	15.50
2	1 678.15	23.75
3	2 316.00	8.00
4	2 061.30	17.00
5	2 207.50	5.00
6	1 708.30	19.00
7	1 784.70	24.00
8	2 575.00	2.50
9	2 357.90	7.50
10	2 277.70	11.00
11	2 165.20	13.00
12	2 399.55	3.75
13	1 779.80	25.00
14	2 336.75	9.75
15	1 765.30	22.00
16	2 053.50	18.00
17	2 414.40	6.00
18	2 200.50	12.50
19	2 654.20	2.00
20	1 753.70	21.50

a) Trace un diagrama de dispersión de los datos. ¿Parecería plausible un modelo de regresión lineal?

b) Encuentre las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen del modelo de regresión lineal simple.

c) Estime la resistencia media al corte de un motor fabricado con una carga propulsora de 20 semanas de antigüedad.

d) Obtenga los valores ajustados $\hat{y}_i$ que corresponden a cada valor observado y_i. Grafique $\hat{y}_i$ contra y_i, y comente la forma que adoptaría esta gráfica si la relación lineal entre la resistencia al corte y la antigüedad fuera perfectamente determinista (sin error). ¿Esta gráfica indica que la antigüedad es una elección razonable de la variable regresora de este modelo?

10-13. Demuestre que en un modelo de regresión lineal simple el punto $(\bar{x}, \bar{y})$ se localiza exactamente sobre la recta de regresión de mínimos cuadrados.

10-14. Considérese el modelo de regresión lineal simple $Y = \beta_0 + \beta_1 x + \varepsilon$. Suponga que el analista quiere usar $z = x - \bar{x}$ como variable de regresión.

a) Utilizando los datos del ejercicio 10-12, construya un diagrama de dispersión de los puntos (x_i, y_i) y otro para los puntos $(z_i = x_i - \bar{x}, y_i)$. Use las dos gráficas para explicar de manera intuitiva cómo se relacionan los dos modelos $Y = \beta_0 + \beta_1 x + \varepsilon$ y $Y = \hat{\beta}_0 + \beta_1^* z + \varepsilon$.

b) Encuentre los estimadores de mínimos cuadrados de β_0^* y β_1^* del modelo $Y = \beta_0^* + \beta_1^* z + \varepsilon$. ¿Cómo se relacionan con los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$?

10-15. Suponga que quiere ajustarse el modelo $Y_i^* = Y_1^* = \beta_0^* + \beta_1^*(x_i - \bar{x}) + \varepsilon$, donde $Y_i^* = Y_i - \bar{Y}$ $(i = 1, 2, ..., n)$. Encuentre los estimadores de mínimos cuadrados de β_0^* y β_1^*. ¿Cómo se relacionan con $\hat{\beta}_0$ y $\hat{\beta}_1$?

10-16. Suponga que quiere ajustarse un modelo de regresión para el que la verdadera recta de regresión pase por el punto (0, 0). El modelo apropiado es $Y = \beta x + \varepsilon$. Suponga que se tienen n pares de datos $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. Encuentre la estimación de mínimos cuadrados de β.

10-17. Utilizando los resultados del ejercicio 10-16, ajuste el modelo $Y = \beta x + \varepsilon$ a los datos de la concentración de cloruros-área de carretera del ejercicio 10-11. Grafique el modelo ajustado en un diagrama de dispersión de los datos y comente si el modelo es apropiado o no.

10-3 PROPIEDADES DE LOS ESTIMADORES DE MÍNIMOS CUADRADOS Y LA ESTIMACIÓN DE σ^2

Es sencillo describir las propiedades estadísticas de los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$. Recuérdese que se ha supuesto que el término del error ε en el modelo $Y = \beta_0 + \beta_1 x + \varepsilon$ es una variable aleatoria con media cero y varianza σ^2. Puesto que los valores de x están fijos, Y es una variable aleatoria con media $\mu_{Y|x} = \beta_0 + \beta_1 x + \varepsilon$ y varianza σ^2. Por lo tanto, los valores de $\hat{\beta}_0$ y $\hat{\beta}_1$ dependen de las y observadas; por tanto, los estimadores de mínimos cuadrados de los coeficientes de regresión pueden considerarse como variables aleatorias. Se investigarán las propiedades del sesgo y de la varianza de los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$.

Se considera primero $\hat{\beta}_1$. Puesto que $\hat{\beta}_1$ es una combinación lineal de las observaciones Y_i, pueden usarse los resultados del capítulo 6 para demostrar que el valor esperado de $\hat{\beta}_1$ es

$$E(\hat{\beta}_1) = \beta_1 \tag{10-12}$$

Por tanto, $\hat{\beta}_1$ es un **estimador insesgado** de la verdadera pendiente β_1.

Considérese ahora la varianza de $\hat{\beta}_1$. Puesto que se ha supuesto que $V(\varepsilon_i) = \sigma^2$, se sigue que $V(Y_i) = \sigma^2$, y puede demostrarse que

$$V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}} \tag{10-13}$$

Aplicando un enfoque similar, puede demostrarse que

$$E(\hat{\beta}_0) = \beta_0 \quad \text{y} \quad V(\hat{\beta}_0) = \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right] \tag{10-14}$$

Por tanto, $\hat{\beta}_0$ es un estimador insesgado de la ordenada al origen β_0. La covarianza de las variables aleatorias $\hat{\beta}_0$ y $\hat{\beta}_1$ es diferente de cero. Es posible demostrar (véase el ejercicio 10-69) que $\text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\sigma^2 \bar{x}/S_{xx}$.

Para hacer inferencias respecto a los coeficientes de regresión β_0 y β_1, es necesario estimar la varianza σ^2 que aparece en las expresiones de $V(\hat{\beta}_0)$ y $V(\hat{\beta}_1)$. El parámetro σ^2, la varianza del término del error ε en el modelo de regresión, refleja la variación aleatoria alrededor de la verdadera recta de regresión.

Se usan los residuales, $e_i = y_i - \hat{y}_i$, para calcular la estimación de σ^2. La suma de los cuadrados de los residuales, o **suma de los cuadrados de los errores** es

$$SS_E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{10-15}$$

Puede demostrarse que el valor esperado de la suma de los cuadrados de los errores SS_E es

$$E(SS_E) = (n-2)\sigma^2 \tag{10-16}$$

Por lo tanto,

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} \tag{10-17}$$

es un estimador insesgado de σ^2.

Una fórmula más conveniente para calcular SS_E puede encontrarse sustituyendo el modelo ajustado $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ en la ecuación 10-15 y simplificando. El resultado es

$$SS_E = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 - \hat{\beta}_1 S_{xy}$$

y si se hace $SS_T \equiv \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2$, entonces SS_E puede escribirse como

$$SS_E = SS_T - \hat{\beta}_1 S_{xy} \tag{10-18}$$

EJEMPLO 10-2

Se encontrará la estimación de la varianza σ^2 utilizando los datos del ejemplo 10-1. Ahora bien, $S_{xy} = 10.18$, $\hat{\beta}_1 = 14.97$, y

$$SS_T = \sum_{i=1}^{20} y_i^2 - \frac{\left(\sum_{i=1}^{20} y_i\right)^2}{20} = 170\,044.53 - \frac{(1\,843.21)^2}{20}$$
$$= 173.37$$

Por lo tanto, al utilizar la ecuación 10-17,

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} = \frac{SS_T - \hat{\beta}_1 S_{xy}}{n-2}$$
$$= \frac{173.37 - (14.97)(10.18)}{20-2} = 1.17$$

Podría utilizarse la estimación de σ^2 ($\hat{\sigma}^2 = 1.17$) en la ecuación 10-14 para obtener una estimación de la varianza de la pendiente, y en la ecuación 10-15 para estimar la varianza de la ordenada al origen. A las raíces cuadradas de los estimadores de las varianzas resultantes se les llama los **errores estándar estimados** de la pendiente y la ordenada al origen, respectivamente.

Definición

En una regresión lineal simple, el **error estándar estimado de la pendiente** es

$$se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

y el **error estándar estimado de la ordenada al origen** es

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}$$

donde $\hat{\sigma}^2$ se calcula con la ecuación 10-17.

La salida del programa de computadora Minitab de la tabla 10-2 reporta los errores estándar de la pendiente y la ordenada al origen, respectivamente. Obsérvese que en dicha tabla la estimación de σ^2 se enlista bajo el encabezado "mean square residual" (cuadrado medio residual) en la entrada de la porción desplegada del análisis de varianza. (Esta información se analizará en la sección 10-5.) Además, debido a que la computadora maneja más cifras decimales en los cálculos que las utilizadas aquí en el cálculo manual, el valor calculado por Minitab para el error cuadrado medio es 1.18, que difiere ligeramente del valor 1.17 obtenido arriba.

10-4 ABUSOS COMUNES DE LA REGRESIÓN

La regresión es de uso generalizado y con frecuencia mal usada; a continuación se mencionan de manera breve varios abusos comunes de la regresión. Deberá ponerse atención al seleccionar las variables con las que se construyen las ecuaciones de regresión así como al determinar la forma del modelo. Es posible desarrollar relaciones estadísticas entre variables que no tienen ninguna relación en un sentido práctico. Por ejemplo, podría intentarse relacionar la resistencia al corte de puntos de soldadura con el número de cajas de papel para computadora utilizadas por el grupo de sistemas de información. Una recta parecería proporcionar incluso un buen ajuste a los datos, pero la relación no tiene bases razonables. Una fuerte asociación observada entre variables no implica necesariamente que existe una relación causal entre las mismas. El diseño de experimentos es la única manera de determinar relaciones causales.

Las relaciones de regresión únicamente son válidas para los valores de la variable de regresión que se encuentran en el rango de los datos originales. La relación lineal que se ha supuesto de manera tentativa puede ser válida en el rango original de x, pero quizá deje de serlo cuando se hace una extrapolación; es decir, cuando se usan valores de x fuera de ese rango. En otras palabras, conforme se rebasa el rango de los valores de x para los que se recabaron los datos, disminuye la certeza acerca de la validez del modelo supuesto. Los modelos de regresión no son válidos necesariamente para fines de extrapolación.

10-5 PRUEBAS DE HIPÓTESIS EN LA REGRESIÓN LINEAL SIMPLE

Un aspecto importante al evaluar la adecuación de un modelo de regresión lineal es la prueba de las hipótesis estadísticas acerca de los parámetros del modelo y la construcción de ciertos intervalos de confianza. En esta sección se discute la prueba de hipótesis en una regresión lineal simple, y en la sección 10-6 se presentan los métodos para construir intervalos de confianza. Para probar hipótesis acerca de la pendiente y la ordenada al origen del modelo de regresión, debe establecerse el supuesto adicional de que el componente del error del modelo, ε, sigue una distribución normal. Por tanto, los supuestos completos son que los errores siguen una distribución normal e independiente con media cero y varianza σ^2, que se abrevia NID $(0, \sigma^2)$.

10-5.1 Uso de las pruebas *t*

Suponga que quieren probarse las hipótesis de que la pendiente es igual a una constante, por ejemplo, $\beta_{1,0}$. Las hipótesis apropiadas son

$$\begin{aligned} H_0 &: \beta_1 = \beta_{1,0} \\ H_1 &: \beta_1 \neq \beta_{1,0} \end{aligned} \tag{10-19}$$

donde se ha supuesto una hipótesis alternativa de dos colas. Puesto que los errores ε_i son NID $(0, \sigma^2)$, se sigue directamente que las observaciones Y_i son NID $(\beta_0 + \beta_1 x_i, \sigma^2)$. Por la ecuación 10-8 se observa que $\hat{\beta}_1$ es una combinación lineal de las observaciones Y_i. Por tanto, $\hat{\beta}_1$ es una combinación lineal de variables aleatorias normales e independientes, y por consiguiente, $\hat{\beta}_1$ es $N(\beta_1, \sigma^2/S_{xx})$, usando las propiedades del sesgo y de la varianza de la pendiente revisadas en la sección 10-3. Además, $(n-2)\hat{\sigma}^2/\sigma^2$ tiene una distribución ji-cuadrada con $n-2$ grados de libertad, y $\hat{\beta}_1$ es independiente de $\hat{\sigma}^2$. Como resultado de dichas propiedades, el estadístico

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\hat{\sigma}^2 / S_{xx}}} \tag{10-20}$$

sigue la distribución t con $n-2$ grados de libertad bajo H_0: $\beta_1 = \beta_{1,0}$. Se rechazaría H_0: $\beta_1 = \beta_{1,0}$ si

$$|t_0| > t_{\alpha/2, n-2} \tag{10-21}$$

donde t_0 se calcula con la ecuación 10-20. El denominador de la ecuación 10-20 es el error estándar de la pendiente, por lo que el estadístico de la prueba podría escribirse como

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{se\left(\hat{\beta}_1\right)}$$

Puede usarse un procedimiento similar para probar hipótesis acerca de la ordenada al origen. Para probar

$$\begin{aligned} H_0 &: \beta_0 = \beta_{0,0} \\ H_1 &: \beta_0 \neq \beta_{0,0} \end{aligned} \tag{10-22}$$

se usaría el estadístico

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}} = \frac{\hat{\beta}_0 - \beta_{0,0}}{se\left(\hat{\beta}_0\right)} \tag{10-23}$$

y la hipótesis nula se rechazaría si el valor calculado de este estadístico de la prueba, t_0, es tal que $|t_0| > t_{\alpha/2,\, n-2}$. Obsérvese que el denominador del estadístico de la prueba de la ecuación 10-23 no es sino el error estándar de la ordenada al origen.

Un caso especial muy importante de las hipótesis de la ecuación 10-19 es

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_1 &: \beta_1 \neq 0 \end{aligned} \tag{10-24}$$

Estas hipótesis están relacionadas con la **significación de la regresión**. No poder rechazar H_0: $\beta_1 = 0$ es equivalente a concluir que no hay una relación lineal entre x y Y. En la figura 10-5 se ilustra esta situación. Obsérvese que esto implica que x es de escaso valor para explicar la variación de Y y que el mejor estimador de Y para cualquier x es $\hat{Y} = \bar{Y}$ (figura 10-5*a*) o bien que la verdadera relación entre x y Y es no lineal (figura 10-5*b*). Alternativamente, si se rechaza H_0: $\beta_1 = 0$,

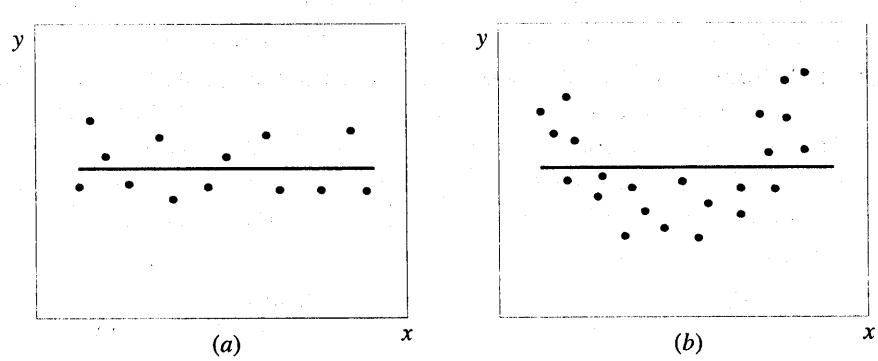


Figura 10-5 La hipótesis $H_0 : \beta_1 = 0$ no se rechaza.

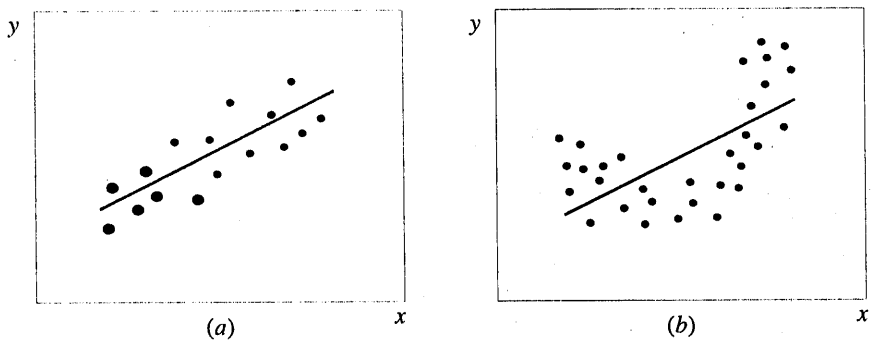


Figura 10-6 La hipótesis H_0: $\beta_1 = 0$ se rechaza.

esto implica que x es de valor para explicar la variabilidad de Y (véase la figura 10-6). El rechazo de H_0: $\beta_1 = 0$ significaría, ya sea que el modelo lineal es adecuado (figura 10-6*a*) o bien que, aun cuando hay un efecto lineal de x, podrían obtenerse mejores resultados con la adición de términos polinómicos en x de orden superior (figura 10-6*b*).

EJEMPLO 10-3

Se probará la significación de la regresión utilizando el modelo para los datos de la pureza del oxígeno del ejemplo 10-1. Las hipótesis son

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

y se usará $\alpha = 0.01$. Por los ejemplos 10-1 y 10-2, se tiene

$$\hat{\beta}_1 = 14.97 \quad n = 20, \quad S_{xx} = 0.68, \quad \hat{\sigma}^2 = 1.17$$

por lo que el estadístico t de la ecuación 10-20 es

$$t_0 = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2 / S_{xx}}} = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} = \frac{14.97}{\sqrt{1.17/0.68}} = 11.41$$

Puesto que el valor de referencia de t es $t_{0.005,\,18} = 2.88$, el valor del estadístico de la prueba se encuentra muy adentro de la región crítica, lo cual implica que H_0: $\beta_1 = 0$ deberá rechazarse. El valor P para esta prueba es $P \simeq 1.13 \times 10^{-9}$. Este valor se obtuvo manualmente con una calculadora.

Como ya se señaló, en la tabla 10-2 se presenta la salida del programa Minitab para este problema. Obsérvese que el valor del estadístico t para la pendiente es 11.35 y que el valor P reportado es $P = 0.00000$. Minitab también reporta el estadístico t para probar la hipótesis H_0: $\beta_0 = 0$. Este estadístico se calcula con la ecuación 10-23, con $\beta_{0,0} = 0$, como $t_0 = 46.62$. Evidentemente, entonces, se rechaza la hipótesis de que la ordenada al origen es cero.

10-5.2 Enfoque del análisis de varianza para probar la significación de una regresión

Puede usarse un método llamado **análisis de varianza** para probar la significación de una regresión. En el procedimiento se hace la partición de la variabilidad total de la variable de respuesta en componentes significativos como base de la prueba. El *análisis de identidad de la varianza* es el siguiente:

$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2+\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \qquad (10\text{-}25)$$

Los dos componentes del miembro derecho de la ecuación 10-25 miden, respectivamente, la cantidad de variabilidad en y_i explicada por la recta de regresión y la variación residual que queda sin explicar por la recta de regresión. A $SS_E=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ suele llamársele la **suma de los cuadrados de errores** y a $SS_R=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$ la **suma de los cuadrados de regresión**. Utilizando símbolos, la ecuación 10-25 puede escribirse como

$$SS_T=SS_R+SS_E \qquad (10\text{-}26)$$

donde $SS_T=\sum_{i=1}^{n}(y_i-\bar{y})^2$ es la **suma total corregida de los cuadrados** de y. En la sección 10-3 se señaló que $SS_E=SS_T+\hat{\beta}_1S_{xy}$ (véase la ecuación 10-18), por lo que como $SS_T=\hat{\beta}_1S_{xy}+SS_E$ se observa que la suma de los cuadrados de regresión de la ecuación 10-26 es

$$SS_R=\hat{\beta}_1S_{xy} \qquad (10\text{-}27)$$

La suma total de los cuadrados SS_T tiene $n-1$ grados de libertad, y SS_R y SS_E tienen 1 y $n-2$ grados de libertad, respectivamente.

Puede demostrarse que $E[SS_E/(n-2)]=\sigma^2$, $E(SS_R)=\sigma^2+\beta_1^2S_{xx}$, y que SS_E/σ^2 y SS_R/σ^2 son variables aleatorias ji-cuadrada independientes con $n-2$ y 1 grados de libertad, respectivamente. Por tanto, si la hipótesis nula H_0: $\beta_1=0$ es verdadera, el estadístico

$$F_0 = \frac{SS_R/1}{SS_E/(n-2)} = \frac{MS_R}{MS_E} \tag{10-28}$$

sigue la distribución $F_{1,n-2}$, y H_0 se rechazaría si $f_0 > f_{\alpha,1,n-2}$. A las cantidades $MS_R = SS_R/1$ y $MS_E = SS_E/(n-2)$ se les llama **cuadrados medios**. En general, un cuadrado medio se calcula siempre dividiendo una suma de cuadrados entre su número de grados de libertad. El procedimiento de prueba suele ordenarse en una **tabla de análisis de varianza**, como la tabla 10-3.

EJEMPLO 10-4

Se usará el enfoque del análisis de varianza para probar la significación de la regresión utilizando los datos de la pureza del oxígeno del ejemplo 10-1. Recuérdese que $SS_T = 173.37$, $\hat{\beta}_1 = 14.97$, $S_{xy} = 10.18$, y $n = 20$. La suma de los cuadrados de regresión es

$$SS_R = \hat{\beta}_1 S_{xy} = (14.97)10.18 = 152.39$$

y la suma de los cuadrados de los errores es

$$\begin{aligned} SS_E &= SS_T - SS_R \\ &= 173.37 - 152.39 \\ &= 20.98 \end{aligned}$$

El análisis de varianza para probar H_0: $\beta_1 = 0$ se resume en la tabla 10-4. El estadístico de la prueba es $f_0 = MS_R/MS_E = 152.39/1.17 = 130.25$, para el que se encuentra que el valor P es $P \simeq 1.13 \times 10^{-9}$, de donde se concluye que β_1 es diferente de cero.

La salida de Minitab de la tabla 10-2 también presenta el análisis de varianza para este problema. Con frecuencia hay diferencias menores en la terminología entre los paquetes de computadora. Por ejemplo, en ocasiones a la suma de cuadrados de regresión se le llama la suma de cuadrados "del modelo", y a la suma de cuadrados de los errores se le llama la suma de los cuadrados "de los residuales". Los resultados numéricos obtenidos con la computadora son muy similares a los que se obtuvieron manualmente arriba.

Tabla 10-3 Análisis de varianza para probar la significación de la regresión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Regresión	$SS_R = \hat{\beta}_1 S_{xy}$	1	MS_R	MS_R/MS_E
Error	$SS_E = SS_T - \hat{\beta}_1 S_{xy}$	$n-2$	MS_E	
Total	SS_T	$n-1$		

Obsérvese que $MS_E = \hat{\sigma}^2$.

Tabla 10-4 Prueba de significación de la regresión, ejemplo 10-4

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Regresión	152.39	1	152.39	130.25	1.13×10^{-9}
Error	20.98	18	1.17		
Total	173.37	19			

Obsérvese que el procedimiento del análisis de varianza para probar la significación de la regresión es equivalente a la prueba t de la sección 10-5.1. Es decir, cualquiera de los dos procedimientos llevará a las mismas conclusiones. Es sencillo demostrar esto empezando con el estadístico de la prueba t de la ecuación 10-20 con $\beta_{1,0} = 0$, por ejemplo

$$T_0 = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2 / S_{xx}}} \tag{10-29}$$

Al elevar al cuadrado ambos miembros de la ecuación 10-29, y al usar el hecho de que $\hat{\sigma}^2 = MS_E$ se obtiene

$$T_0^2 = \frac{\hat{\beta}_1^2 S_{xx}}{MS_E} = \frac{\hat{\beta}_1 S_{xY}}{MS_E} = \frac{MS_R}{MS_E} \tag{10-30}$$

Obsérvese que T_0^2 de la ecuación 10-30 es idéntica a F_0 de la ecuación 10-28. Se cumple, en general, que el cuadrado de una variable aleatoria t con v grados de libertad es una variable aleatoria F, con 1 y v grados de libertad en el numerador y el denominador, respectivamente. Por tanto, la prueba utilizando T_0 es equivalente a la prueba basada en F_0. Sin embargo, cabe observar que la prueba t es un tanto más flexible por cuanto permitiría hacer la prueba contra una hipótesis alternativa de una cola, en tanto que la prueba F está restringida a una hipótesis alternativa de dos colas.

EJERCICIOS DE LA SECCIÓN 10-5

10-18. Considérense los datos del ejercicio 10-1 para x = resistencia a la compresión y y = permeabilidad intrínseca del concreto.

a) Pruebe la significación de la regresión utilizando $\alpha = 0.05$. Encuentre el valor P para esta prueba. ¿Puede concluirse que el modelo especifica una relación lineal útil entre estas dos variables?

b) Estime σ^2 y la desviación estándar de $\hat{\beta}_1$.

c) ¿Cuál es el error estándar de la ordenada al origen en este modelo?

10-19. Considérense los datos del ejercicio 10-2 para x = temperatura superficial de una carretera y y = deflexión del pavimento.

a) Pruebe la significación de la regresión utilizando $\alpha = 0.05$. Encuentre el valor P para esta prueba. ¿Qué conclusiones pueden sacarse?

b) Estime σ^2.

c) Estime los errores estándar de la pendiente y la ordenada al origen.

10-20. Considérense los datos de la Liga Nacional de Futbol Americano (NFL) del ejercicio 10-4.

a) Pruebe la significación de la regresión utilizando $\alpha = 0.01$. Encuentre el valor P para esta prueba. ¿Qué conclusiones pueden sacarse?

b) Estime σ^2.

c) Estime los errores estándar de la pendiente y la ordenada al origen.

d) Pruebe (usando $\alpha = 0.01$) H_0: $\beta_1 = -0.01$ contra H_1: $\beta_1 \neq -0.01$. ¿Estaría usted de acuerdo en que esta es una prueba de la afirmación de que si un equipo puede reducir 100 yardas el yardaje corrido en el campo del oponente ganará un juego más?

10-21. Considérense los datos del ejercicio 10-5 para y = precio de venta y x = impuestos pagados.

a) Pruebe H_0: $\beta_1 = 0$ usando la prueba t; utilice $\alpha = 0.05$.

b) Pruebe H_0: $\beta_1 = 0$ usando el análisis de varianza con $\alpha = 0.05$. Comente la relación de esta prueba con la prueba del inciso a) anterior.

c) Estime σ^2, y estime los errores estándar de la pendiente y la ordenada al origen.

d) Pruebe la hipótesis de que $\beta_0 = 0$.

10-22. Considérense los datos del ejercicio 10-6 para y = consumo de vapor y x = temperatura promedio.

a) Pruebe la significación de la regresión usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba? Enuncie las conclusiones que resultan de esta prueba.

b) Estime σ^2.

c) Estime los errores estándar de la pendiente y la ordenada al origen.

d) Pruebe la hipótesis H_0: $\beta_1 = 10$ contra H_1: $\beta_1 \neq 10$ usando $\alpha = 0.01$. Encuentre el valor P para esta prueba.

e) Pruebe H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.01$. Encuentre el valor P para esta prueba y saque conclusiones.

10-23. En el ejercicio 10-7 se dieron 20 observaciones para y = rendimiento de combustible en carretera y x = desplazamiento del motor.

a) Pruebe la significación de la regresión usando $\alpha = 0.01$. Encuentre el valor P para esta prueba. ¿A qué conclusiones puede llegarse?

b) Estime σ y los errores estándar de la pendiente y la ordenada al origen.

c) Pruebe H_0: $\beta_1 = -0.05$ contra H_1: $\beta_1 < -0.05$ usando $\alpha = 0.01$ y saque conclusiones. ¿Cuál es el valor P para esta prueba?

d) Pruebe la hipótesis H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?

10-24. En el ejercicio 10-8 se dieron 13 observaciones sobre y = concentración del licor verde Na_2S y x = producción en una fábrica de papel.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. Encuentre el valor P para esta prueba.

b) Estime σ y los errores estándar de la pendiente y la ordenada al origen.

c) Pruebe H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

10-25. En el ejercicio 10-9 se presentaron los datos para y = aumento de la presión sanguínea y x = nivel de presión del sonido.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. Encuentre el valor P para esta prueba.

b) Estime σ^2.

c) Estime los errores estándar de la pendiente y la ordenada al origen.

d) Pruebe H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.05$. Encuentre el valor P para esta prueba.

10-26. En el ejercicio 10-11 se presentaron los datos para y = concentración de cloruros en corrientes superficiales y x = área de carretera.

a) Pruebe la hipótesis H_0: $\beta_1 = 0$ contra H_1: $\beta_1 \neq 0$ usando el procedimiento del análisis de varianza con $\alpha = 0.01$.

b) Encuentre el valor P para la prueba del inciso a).

c) Estime σ^2 y los errores estándar de $\hat{\beta}_1$ y $\hat{\beta}_0$.

d) Pruebe H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.01$. ¿Qué conclusiones pueden

sacarse? ¿El modelo parecería un mejor ajuste a los datos si se eliminara la ordenada al origen?

10-27. Con referencia al ejercicio 10-12, donde se dan 20 observaciones para y = resistencia al corte de una carga propulsora y x = antigüedad de la carga propulsora.

a) Pruebe la significación de la regresión con $\alpha = 0.01$. Encuentre el valor P para esta prueba.

b) Estime σ^2 y los errores estándar de $\hat{\beta}_0$ y $\hat{\beta}_1$.

c) Pruebe H_0: $\beta_1 = -30$ contra H_1: $\beta_1 \neq -30$ usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?

d) Pruebe H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?

e) Pruebe H_0: $\beta_0 = 2\,500$ contra H_1: $\beta_0 > 2\,500$ usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?

10-28. Suponga que cada valor de x_i se multiplica por una constante positiva a, y que cada valor de y_i se multiplica por otra constante positiva b. Demuestre que no cambia el valor del estadístico t para probar H_0: $\beta_1 = 0$ contra H_1: $\beta_1 \neq 0$.

10-29. Considérese el modelo sin ordenada al origen, $Y = \beta x + \varepsilon$ con NID $(0, \sigma^2)$ de las ε. La estimación de σ^2 es $s^2 = \sum_{i=1}^{n} \left(y_i - \hat{\beta} x_i\right)^2 / (n-1)$ y $V(\hat{\beta}) = \sigma^2 / \sum_{i=1}^{n} x_i^2$.

a) Idee un estadístico de la prueba para H_0: $\beta = 0$ contra H_1: $\beta \neq 0$.

b) Aplique la prueba del inciso a) al modelo del ejercicio 10-17.

10-30. La probabilidad del error tipo II para la prueba t de H_0: $\beta_1 = \beta_{1,0}$ puede calcularse en forma similar a las pruebas t del capítulo 9. Si el verdadero valor de β_1 es $\hat{\beta}_1'$, el valor

$$d = \frac{\left|\beta_{1,0} - \beta_1'\right|}{\sigma\sqrt{\dfrac{n-1}{S_{xx}}}}$$

se calcula y se usa como factor de la escala horizontal en las curvas de operación característica para la prueba t (gráficas VI*e* a VI*h* del apéndice), y la probabilidad del error tipo II se lee en la escala vertical usando la curva para $n - 2$ grados de libertad. Aplique este procedimiento a los datos del futbol americano del ejercicio 10-4 usando $\sigma = 2.4$ y $\beta_1' = -0.005$, donde las hipótesis son H_0: $\beta_1 = -0.01$ contra H_1: $\beta_1 \neq -0.01$.

10-6 INTERVALOS DE CONFIANZA

10-6.1 Intervalos de confianza para la pendiente y la ordenada al origen

Además de las estimaciones puntuales de la pendiente y la ordenada al origen, es posible obtener estimaciones de los intervalos de confianza para estos parámetros. La anchura de estos intervalos de confianza es una medida de la calidad global de la recta de regresión. Si los términos del error, ε_i, del modelo de regresión tienen una distribución normal e independiente, entonces

$$\left(\hat{\beta}_1 - \beta_1\right) / \sqrt{\sigma^2 / S_{xx}} \quad \text{y} \quad \left(\hat{\beta}_0 - \beta_0\right) / \sqrt{\sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}$$

tienen ambos una distribución igual a la de una variable aleatoria t con $n - 2$ grados de libertad. Esto lleva a la siguiente definición de los intervalos de confianza del $100(1 - \alpha)$ por ciento para la pendiente y la ordenada al origen.

Definición

Bajo el supuesto de que las observaciones tienen una distribución normal e independiente, un **intervalo de confianza del 100(1 – α) por ciento para la pendiente** β_1 en una regresión lineal simple es

$$\hat{\beta}_1 - t_{\alpha/2,\,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2,\,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \qquad (10\text{-}31)$$

De manera similar, un **intervalo de confianza del 100(1 – α) por ciento para la ordenada al origen** β_0 es

$$\hat{\beta}_0 - t_{\alpha/2,\,n-2}\sqrt{\sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]} \le \beta_0 \le \hat{\beta}_0 + t_{\alpha/2,\,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]} \qquad (10\text{-}32)$$

EJEMPLO 10-5

Se encontrará un intervalo de confianza de 95% para la pendiente de la recta de regresión usando los datos del ejemplo 10-1. Recuérdese que $\hat{\beta}_1 = 14.97$, $S_{xx} = 0.68$ y $\hat{\sigma}^2 = 1.17$ (véase la tabla 10-2). Entonces, por la ecuación 10-31 se encuentra

$$\hat{\beta}_1 - t_{0.025,\,18}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \le \beta_1 \le \hat{\beta}_1 + t_{0.025,\,18}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

o bien

$$14.97 - 2.101\sqrt{\frac{1.17}{0.68}} \le \beta_1 \le 14.97 + 2.101\sqrt{\frac{1.17}{0.68}}$$

Y al simplificar esta expresión se obtiene

$$12.21 \le \beta_1 \le 17.73$$

10-6.2 Intervalo de confianza para la respuesta media

Puede construirse un intervalo de confianza a partir de la respuesta media en un valor especificado de x, por ejemplo, x_0. Éste es un intervalo de confianza alrededor de $E(Y \mid x_0) = \mu_{Y|x_0}$, al que suele llamársele intervalo de confianza alrededor de la recta de regresión. Puesto que $E(Y \mid x_0) = \mu_{Y|x_0} = \beta_0 + \beta_1 x_0$, puede obtenerse una estimación puntual de $\mu_{Y|x_0}$ a partir del modelo ajustado como

$$\hat{\mu}_{Y|x_0} = \hat{\beta}_0 + \hat{\beta}_1 x_0$$

Entonces $\hat{\mu}_{Y|x_0}$ es un estimador puntual insesgado de $\mu_{Y|x_0}$, ya que $\hat{\beta}_0$ y $\hat{\beta}_1$ son estimadores insesgados de β_0 y β_1. La varianza de $\hat{\mu}_{Y|x_0}$ es

$$V\left(\hat{\mu}_{Y|x_0}\right) = \sigma^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]$$

Este último resultado se sigue del hecho de que $\operatorname{cov}\left(\bar{Y}, \hat{\beta}_1\right) = 0$. (Véase el ejercicio 10-71.) Además, $\hat{\mu}_{Y|x_0}$ tiene una distribución normal, ya que $\hat{\beta}_1$ y $\hat{\beta}_0$ tienen una distribución normal, y si se usa $\hat{\sigma}^2$ como estimación de σ^2, es sencillo demostrar que

$$\frac{\hat{\mu}_{Y|x_0} - \mu_{Y|x_0}}{\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}}$$

tiene una distribución t con $n - 2$ grados de libertad. Esto lleva a la siguiente definición del intervalo de confianza.

Definición

Un **intervalo de confianza del 100(1 – α) por ciento alrededor de la respuesta media** en el valor de $x = x_0$, simbolizado $\hat{\mu}_{Y|x_0}$, está dado por

$$\hat{\mu}_{Y|x_0} - t_{\alpha/2,\, n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}$$

$$\leq \mu_{Y|x_0} \leq \hat{\mu}_{Y|x_0} + t_{\alpha/2,\, n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]} \qquad (10\text{-}33)$$

donde $\hat{\mu}_{Y|x_0} = \hat{\beta}_0 + \hat{\beta}_1 x_0$ se calcula a partir del modelo de regresión ajustado.

Obsérvese que la anchura del intervalo de confianza para $\mu_{Y|x_0}$ es una función del valor especificado para x_0. La anchura del intervalo es un mínimo para $x_0 = \bar{x}$ y se ensancha conforme $|x_0 - \bar{x}|$ se incrementa.

EJEMPLO 10-6

Se construirá un intervalo de confianza de 95% alrededor de la respuesta media para los datos del ejemplo 10-1. El modelo ajustado es $\hat{\mu}_{Y|x_0} = 74.20 + 14.97x_0$, y el intervalo de confianza de 95% para $\mu_{Y|x_0}$ se encuentra con la ecuación 10-33 como

$$\hat{\mu}_{Y|x_0} \pm 2.101\sqrt{1.17\left[\frac{1}{20}+\frac{(x_0-1.20)^2}{0.68}\right]}$$

Suponga que hay interés en predecir la pureza media del oxígeno cuando $x_0 = 1.00\%$. Entonces

$$\hat{\mu}_{Y|x_0} = 74.20 + 14.97(1.00) = 89.17$$

y el intervalo de confianza de 95% es

$$\left[89.17 \pm 2.101\sqrt{1.17\left[\frac{1}{20}+\frac{(1.00-1.20)^2}{0.68}\right]}\right]$$

o bien

$$89.17 \pm 0.75$$

Por lo tanto, el intervalo de confianza de 95% para $\mu_{Y|1.00}$ es

$$88.42 \le \mu_{Y|1.00} \le 89.92$$

Al repetir estos cálculos para varios valores diferentes de x_0 pueden obtenerse los límites de confianza para cada uno de los valores correspondientes de $\mu_{Y|x_0}$. En la figura 10-7 se muestra el diagrama de dispersión con el modelo ajustado y los límites de confianza de 95% correspondientes graficados como las líneas superior e inferior. El nivel de confianza de 95% sólo se aplica al intervalo obtenido para un valor de x, y no al conjunto completo de niveles x. Obsérvese que la anchura del intervalo de confianza para $\mu_{Y|x_0}$ se incrementa conforme $|x_0 - \bar{x}|$ se incrementa.

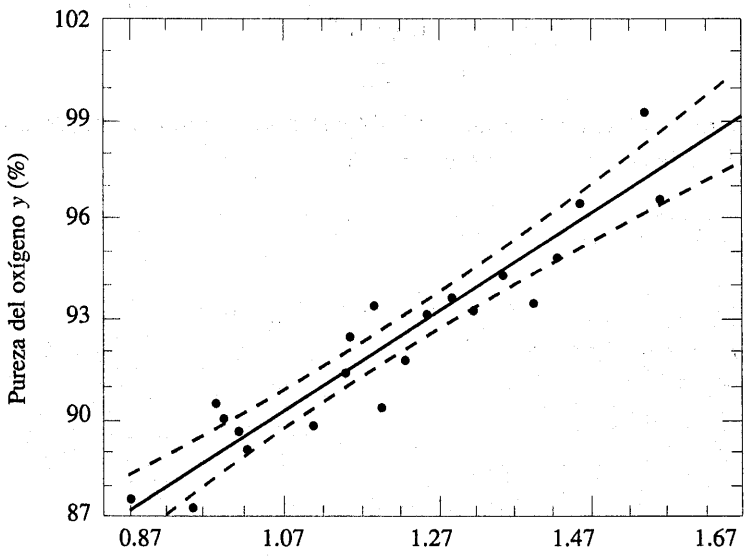


Figura 10-7 Diagrama de dispersión de los datos de la pureza del oxígeno del ejemplo 10-1 con la recta de regresión ajustada y los límites de confianza de 95% para $\mu_{Y|x_0}$.

10-7 PREDICCIÓN DE NUEVAS OBSERVACIONES

Una aplicación importante de los modelos de regresión es predecir observaciones nuevas o futuras Y que correspondan a un nivel especificado de la variable de regresión x. Si x_0 es el valor de la variable de regresión de interés, entonces

$$\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0 \tag{10-34}$$

es el estimador puntual del valor nuevo o futuro de la respuesta Y_0.

Se considera ahora cómo obtener la estimación de un intervalo para esta observación futura Y_0. Esta nueva observación es independiente de las observaciones utilizadas para desarrollar el modelo de regresión. Por lo tanto, el intervalo de confianza para $\mu_{Y|x_0}$ de la ecuación 10-33 no es apropiado, ya que se basa tan sólo en los datos que se utilizaron para hacer el ajuste del modelo de regresión. El intervalo de confianza alrededor de $\mu_{Y|x_0}$ se refiere a la verdadera respuesta media para $x = x_0$ (es decir, a un parámetro poblacional), no a observaciones futuras.

Sea Y_0 la observación futura para $x = x_0$, y sea que $\hat{Y}_0$ dada por la ecuación 10-34 sea el estimador de Y_0. Obsérvese que el error de la predicción

$$\Psi = Y_0 - \hat{Y}_0$$

es una variable aleatoria con una distribución normal con media cero y varianza

$$\begin{aligned} V(\Psi) &= V(Y_0 - \hat{Y}_0) \\ &= \sigma^2\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right] \end{aligned}$$

debido a que Y_0 es independiente de $\hat{Y}_0$. Si se usa $\hat{\sigma}^2$ para estimar σ^2, puede demostrarse que

$$\frac{Y_0 - \hat{Y}_0}{\sqrt{\hat{\sigma}^2\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}}$$

tiene una distribución t con $n - 2$ grados de libertad. A partir de lo anterior es posible desarrollar la siguiente definición de **intervalo de predicción**.

Definición

El **intervalo de predicción del 100(1 – α) por ciento para una observación futura** y_0 para el valor x_0 está dado por

$$\hat{y}_0 - t_{\alpha/2,\, n-2}\sqrt{\hat{\sigma}^2\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]} \le y_0 \le \hat{y}_0 + t_{\alpha/2,\, n-2}\sqrt{\hat{\sigma}^2\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]} \tag{10-35}$$

El valor $\hat{y}_0$ se calcula a partir del modelo de regresión $\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$.

Obsérvese que el intervalo de predicción tiene una anchura mínima para $x_0 = \bar{x}$ y que se ensancha conforme $|x_0 - \bar{x}|$ se incrementa. Al comparar la ecuación 10-35 con la ecuación 10-33, se observa que el intervalo de predicción en el punto x_0, siempre tiene una anchura mayor que el intervalo de confianza para x_0. Esto se debe a que el intervalo de predicción depende tanto del error del modelo ajustado como del error asociado con las observaciones futuras.

EJEMPLO 10-7

Para ilustrar la construcción de un intervalo de predicción, suponga que se usan los datos del ejemplo 10-1 y se encuentra un intervalo de predicción de 95% para la siguiente observación de la pureza del oxígeno para $x_0 = 1.00\%$. Usando la ecuación 10-35 y recordando por el ejemplo 10-6 que $\hat{y}_0 = 89.17$, se encuentra que el intervalo de predicción es

$$89.17 - 2.101\sqrt{1.17\left[1 + \frac{1}{20} + \frac{(1.00-1.20)^2}{0.68}\right]}$$

$$\leq y_0 \leq 89.17 + 2.101\sqrt{1.17\left[1 + \frac{1}{20} + \frac{(1.00-1.20)^2}{0.68}\right]}$$

y al simplificar esta expresión se obtiene

$$86.78 \leq y_0 \leq 91.56$$

Al repetir los cálculos anteriores en diferentes niveles de x_0 pueden obtenerse los intervalos de predicción de 95% que se ilustran gráficamente en la figura 10-8, como las líneas inferiores y superiores alrededor del modelo de regresión ajustado. Obsérvese que en esta gráfica también se muestran los límites de confianza de 95% para $\mu_{Y|x_0}$, los cuales se calcularon en el ejemplo 10-6. En esta gráfica se observa que los límites de predicción son siempre más grandes que los límites de confianza.

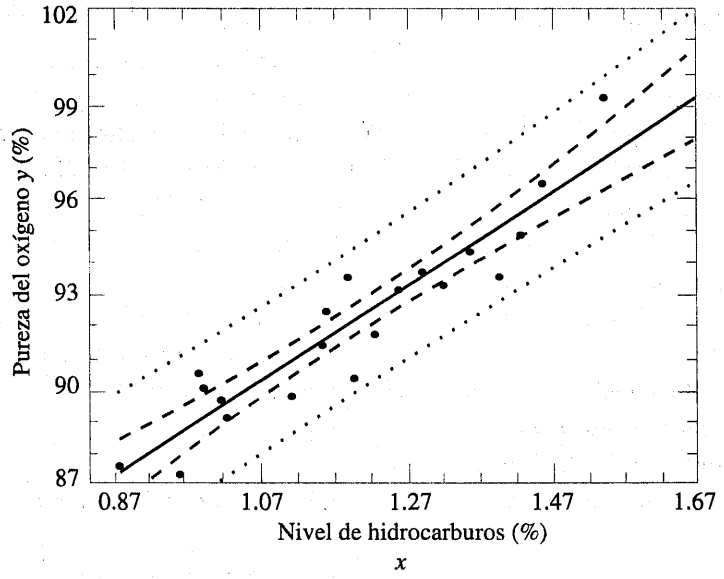


Figura 10-8 Diagrama de dispersión de los datos de la pureza del oxígeno del ejemplo 10-1 con la recta de regresión ajustada, los límites de predicción de 95% (líneas exteriores) y los límites de confianza de 95% para $\mu_{Y|x_0}$.

EJERCICIOS DE LAS SECCIONES 10-6 Y 10-7

10-31. Remítase a los datos del ejercicio 10-1 para y = permeabilidad intrínseca del concreto y x = resistencia a la compresión. Encuentre un intervalo de confianza de 95% para:

a) la pendiente.
b) la ordenada al origen.
c) la permeabilidad media cuando $x = 2.5$.
d) Encuentre un intervalo de predicción de 95% para la permeabilidad cuando $x = 2.5$. Explique por qué este intervalo tiene una anchura mayor que el intervalo del inciso c).

10-32. En el ejercicio 10-2 se presentaron los datos de la temperatura superficial de una carretera x y la deflexión del pavimento y. Encuentre un intervalo de confianza de 99% para:

a) la pendiente.
b) la ordenada al origen.
c) la deflexión media cuando la temperatura es $x = 85°F$.
d) Encuentre un intervalo de predicción de 99% para la deflexión del pavimento cuando la temperatura es 90°F.

10-33. En el ejercicio 10-4 se presentaron los datos sobre el número de juegos ganados por los equipos de la NFL en 1976. Encuentre un intervalo de confianza de 95% para:

a) la pendiente.
b) la ordenada al origen.
c) el número promedio de juegos ganados cuando las yardas corridas en el campo de los oponentes es limitado a $x = 1\ 800$.
d) Encuentre un intervalo de predicción para el número de juegos ganados cuando las yardas corridas en el campo de los oponentes son $x = 1\ 800$.

10-34. Remítase a los datos del ejercicio 10-5 para y = precio de venta de una casa y x = impuestos pagados. Encuentre un intervalo de confianza de 95% para los siguientes datos:

a) β_1
b) β_0
c) el precio de venta promedio cuando los impuestos pagados son $x = 7.50$
d) Calcule el intervalo de predicción de 95% para el precio de venta cuando los impuestos pagados son $x = 7.50$.

10-35. En el ejercicio 10-6 se presentaron los datos para y = consumo de vapor y x = temperatura mensual promedio.

a) Encuentre un intervalo de confianza de 99% para β_1.
b) Encuentre un intervalo de confianza de 99% para β_0.
c) Encuentre un intervalo de confianza de 95% para el consumo promedio de vapor cuando la temperatura promedio es 55°F.
d) Encuentre un intervalo de predicción de 95% para el consumo de vapor cuando la temperatura es 55°F. Explique por qué este intervalo tiene una anchura mayor que la del intervalo del inciso c).

10-36. En el ejercicio 10-7 se presentó el rendimiento de gasolina de 20 automóviles, junto con información acerca del cilindraje del motor. Encuentre un intervalo de confianza de 95% para cada uno de los siguientes datos:

a) la pendiente
b) la ordenada al origen
c) las millas promedio por galón de gasolina en carretera cuando el cilindraje del motor es $x = 150$ pulg^3.
d) Construya un intervalo de predicción de 95% para el rendimiento de gasolina en carretera cuando el cilindraje del motor es $x = 150$ pulg^3.

10-37. Considérense los datos del ejercicio 10-8 para y = concentración del licor verde Na_2S y x = producción en una fábrica de papel. Encuentre un intervalo de confianza de 99% para lo siguiente:

a) β_1
b) β_0
c) la concentración media de Na_2S cuando la producción es $x = 910$ toneladas/día.
d) Encuentre un intervalo de predicción de 99% para la concentración de Na_2S cuando $x = 910$ toneladas/día.

10-38. En el ejercicio 10-9 se presentaron los datos para y = aumento de la presión sanguínea y x = nivel de presión del sonido. Encuentre un intervalo de confianza de 95% para lo siguiente:

a) β_1

b) β_0

c) el aumento promedio de la presión sanguínea cuando el nivel de presión del sonido es 85 dB.

d) Encuentre un intervalo de predicción de 95% para el aumento de la presión sanguínea cuando el nivel de presión del sonido es 85 dB.

10-39. Remítase a los datos del ejercicio 10-10 para y = volumen del desgaste del acero dulce y x = viscosidad del aceite. Encuentre un intervalo de confianza de 95% para lo siguiente:

a) la ordenada al origen.

b) la pendiente.

c) el desgaste medio cuando la viscosidad del aceite es $x = 30$.

10-40. En el ejercicio 10-11 se presentaron los datos para la concentración de cloruros y y el área de carretera x en vertientes de la parte central de Rhode Island. Encuentre un intervalo de confianza de 99% para lo siguiente:

a) β_1

b) β_0

c) la concentración media de cloruros cuando el área de carretera es $x = 1.0\%$.

d) Encuentre un intervalo de predicción de 99% para la concentración de cloruros cuando el área de carretera es $x = 1.0\%$.

10-41. Remítase a los datos del ejercicio 10-12 para la resistencia al corte de un motor de turbina y y la antigüedad de la carga propulsora x. Encuentre un intervalo de confianza de 95% para lo siguiente:

a) la pendiente β_1.

b) la ordenada al origen β_0.

c) la resistencia al corte medio cuando la antigüedad es $x = 20$ semanas.

d) Encuentre un intervalo de predicción de 95% para la resistencia al corte cuando la antigüedad es $x = 20$ semanas.

10-8 EVALUANDO LA ADECUACIÓN DEL MODELO DE REGRESIÓN

El ajuste de un modelo de regresión requiere varios supuestos. La estimación de los parámetros del modelo requiere el supuesto de que los errores son variables aleatorias no correlacionadas con media cero y varianza constante. Las pruebas de hipótesis y la estimación de intervalos requieren que los errores tengan una distribución normal. Además, se supone que el orden del modelo es correcto; es decir, si se ajusta un modelo de regresión lineal simple, entonces se está suponiendo que el fenómeno se comporta en realidad de manera lineal o de primer orden.

El analista deberá considerar siempre la validación de estos supuestos, deberá ser precavido y llevar a cabo análisis a fin de examinar la adecuación del modelo que se ha propuesto de manera tentativa. En esta sección se analizan los métodos convenientes a este respecto.

10-8.1 Análisis de residuales

Los residuales de un modelo de regresión son $e_i = y_i - \hat{y}_i$, $i = 1, 2, ..., n$, donde y_i es una observación real y $\hat{y}_i$ es el valor ajustado correspondiente con el modelo de regresión. El análisis de los residuales con frecuencia resulta de utilidad para verificar el supuesto de que los errores siguen una distribución normal aproximada con varianza constante, así como para determinar si sería conveniente incluir términos adicionales en el modelo.

Como verificación aproximada de la normalidad, el experimentador puede construir un histograma de frecuencias de los residuales o una **gráfica de probabilidad normal de los**

residuales. Muchos programas de computadora generan una gráfica de probabilidad normal de los residuales, y dado que el tamaño de las muestras en una regresión con frecuencia es demasiado pequeño para que un histograma sea significativo, se prefiere el método de las gráficas de probabilidad normal. Determinar la ausencia de normalidad en estas gráficas requiere criterio. (Remítase al tema del método del "lápiz grueso" del capítulo 5.)

También puede hacerse la *estandarización* de los residuales calculando $d_i = e_i / \sqrt{\hat{\sigma}^2}$, $i = 1, 2, ..., n$. Si los errores siguen una distribución normal, entonces aproximadamente 95% de los residuales estandarizados deberán estar incluidos en el intervalo (–2, +2). Los residuales que se aparten considerablemente de este intervalo pueden indicar la presencia de un **punto atípico**, es decir, una observación que no es típica del resto de los datos. Se han propuesto varias reglas para descartar los puntos atípicos. Sin embargo, los puntos atípicos en ocasiones proporcionan información importante acerca de circunstancias excepcionales que son de interés para los experimentadores y, por tanto, no deberán descartarse. Para una discusión detallada de los puntos atípicos, véase Montgomery y Peck.

Con frecuencia resulta conveniente graficar los residuales 1) en secuencias de tiempo (si se conoce), 2) contra las $\hat{y}_i$, y 3) contra la variable independiente x. Estas gráficas por lo general tendrán una forma parecida a la de los cuatro patrones generales que se muestran en la figura 10-9. El patrón *a*) de la figura 10-9 representa la situación ideal, mientras que los patrones *b*), *c*) y *d*) representan anomalías. Si la forma de los residuales es como en *b*), la varianza de las observaciones puede incrementarse con el tiempo o con la magnitud de y_i o x_i. Es común usar la transformación de datos de la respuesta y a fin de eliminar este problema. Entre las transformaciones para estabilizar la varianza más usadas se encuentran la utilización de $\sqrt{y}$, $ln\ y$, o $1/y$ como variable de respuesta. Véase Montgomery y Peck para mayores detalles en cuanto a los

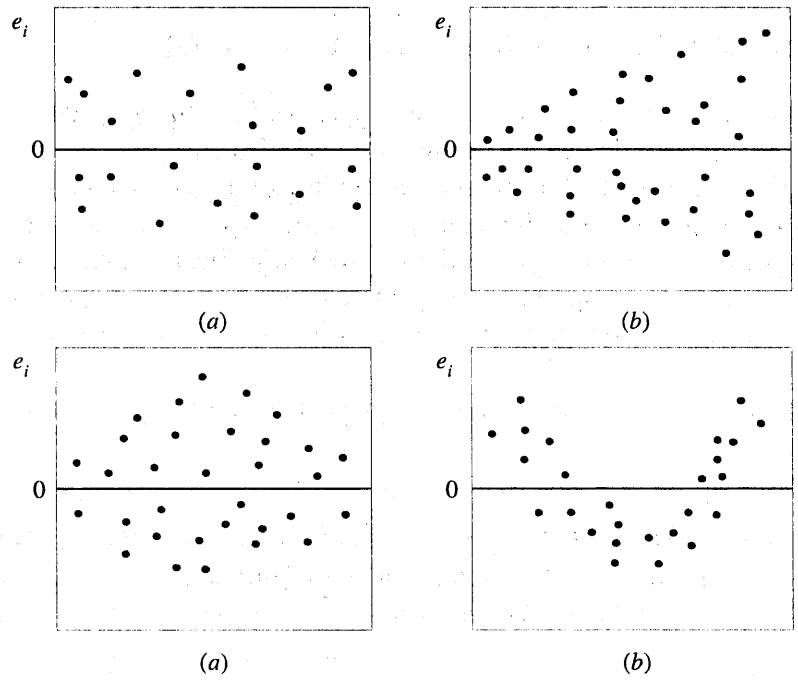


Figura 10-9 Patrones de las gráficas de los residuales: *a*) satisfactoria, *b*) embudo, *c*) arco doble, *d*) no lineal. (Adaptación de Montgomery y Peck.)

métodos para seleccionar la transformación apropiada. Si la forma de la gráfica de los residuales contra el tiempo es como en *b*), entonces la varianza de las observaciones se incrementa con el tiempo. Las gráficas de los residuales contra $\hat{y}_i$ y x_i cuya forma es como en *c*) también indican la desigualdad de la varianza. Las gráficas de los residuales cuya forma es como en *d*) indican que el modelo no es adecuado; es decir, deberán agregarse términos de órdenes superiores al modelo, deberá considerarse una transformación de la variable x o de la variable y (o de ambas), o bien deberán considerarse otros regresores (variables de regresión).

EJEMPLO 10-8

El modelo de regresión para los datos de la pureza del oxígeno del ejemplo 10-1 es $\hat{y} = 74.20 + 14.97x$. En la tabla 10-5 se presentan los valores observados y los valores predichos de y para cada valor de x de este conjunto de datos, junto con el residual correspondiente. Estos valores se calcularon usando el programa Minitab y muestran el número típico de cifras decimales de una salida de computadora. En la figura 10-10 se ilustra una gráfica de probabilidad normal de los residuales. Puesto que en la figura los residuales caen aproximadamente sobre una recta, se concluye que no hay una desviación importante de la normalidad. En la figura 10-11 los residuales también se grafican contra los valores predichos $\hat{y}_i$, y contra los niveles de hidrocarburos x_i en la figura 10-12. Estas gráficas no indican alguna falta de adecuación seria del modelo.

Tabla 10-5 Datos de la pureza del oxígeno del ejemplo 10-1, valores predichos y residuales

	Nivel de hidrocarburos, x	Pureza del oxígeno, y	Valor predicho, $\hat{y}$	Residual $e = y - \hat{y}$
1	0.99	90.01	89.069009	0.940991
2	1.02	89.05	89.518136	–0.468136
3	1.15	91.43	91.464353	–0.034353
4	1.29	93.74	93.560279	0.179721
5	1.46	96.73	96.105332	0.624668
6	1.36	94.45	94.608242	–0.158242
7	0.87	87.59	87.272501	0.317499
8	1.23	91.77	92.662025	–0.892025
9	1.55	99.42	97.452713	1.967287
10	1.40	93.65	95.207078	–1.557078
11	1.19	93.54	92.063189	1.476811
12	1.15	92.52	91.614062	0.905938
13	0.98	90.56	88.919300	1.640700
14	1.01	89.54	89.368427	0.171573
15	1.11	89.85	90.865517	–1.015517
16	1.20	90.39	92.212898	–1.822898
17	1.26	93.25	93.111152	0.138848
18	1.32	93.41	94.009406	–0.599406
19	1.43	94.98	95.656205	–0.676205
20	0.95	87.33	88.470173	–1.140173

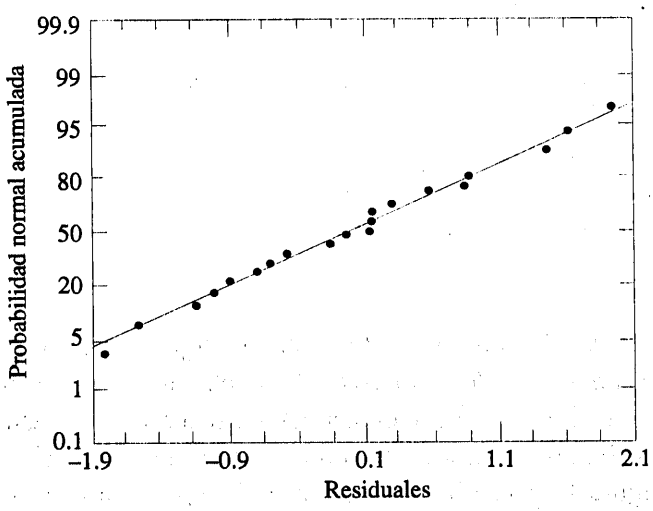


Figura 10-10 Gráfica de probabilidad normal de residuales, ejemplo 10-8.

10-8.2 Coeficiente de determinación (R^2)

A la cantidad

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \tag{10-36}$$

se le llama el **coeficiente de determinación** y se usa con frecuencia para juzgar la adecuación de un modelo de regresión. Más adelante se verá que en el caso en que X y Y son variables aleatorias con una distribución conjunta, R^2 es el cuadrado del coeficiente de correlación entre X y Y. Por el análisis de identidad de la varianza de las ecuaciones 10-25 y 10-26, $0 \leq R^2 \leq 1$. Con

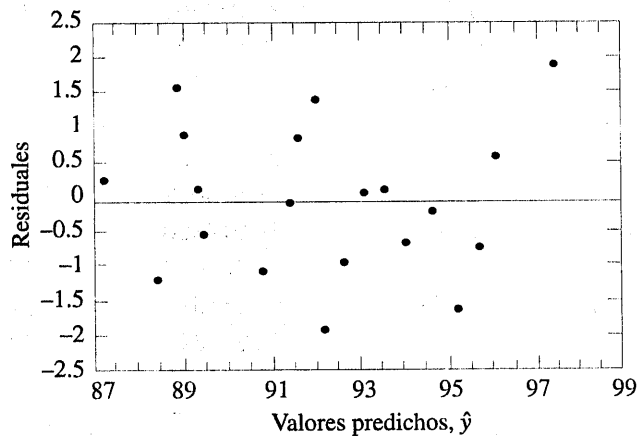


Figura 10-11 Gráfica de los residuales contra la pureza del oxígeno predicha $\hat{y}$, ejemplo 10-8.

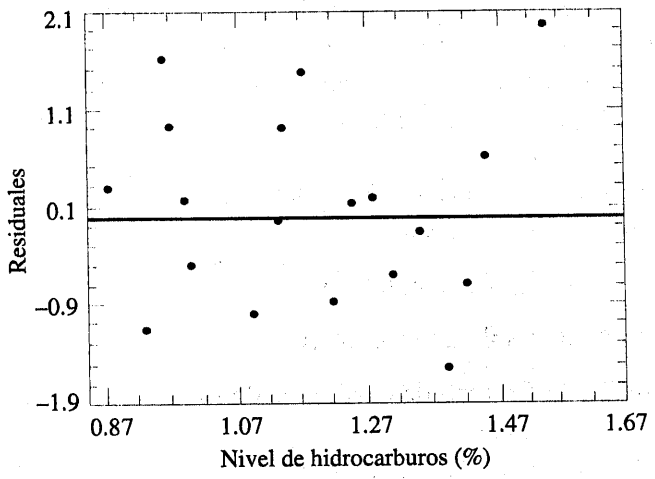


Figura 10-12 Gráfica de los residuales contra el nivel de hidrocarburos x, ejemplo 10-8.

frecuencia se hace referencia a R^2 como la cantidad de la variabilidad de los datos que está explicada o considerada por el modelo de regresión. En el modelo de regresión para la pureza del oxígeno, se tiene $R^2 = SS_R/SS_T = 152.39/173.37 = 0.8790$; es decir, el modelo explica 87.90% de la variabilidad de los datos.

El estadístico R^2 deberá usarse con cuidado, ya que siempre es posible hacer R^2 igual a la unidad mediante la simple adición de términos suficientes al modelo. Por ejemplo, puede obtenerse un ajuste "perfecto" para n datos o puntos observados con un polinomio de grado $n - 1$. Además, R^2 siempre se incrementará si se agrega una variable al modelo, aunque esto no implica necesariamente que el nuevo modelo sea superior al original. A menos que la suma de cuadrados de los errores del nuevo modelo tenga una reducción por una cantidad igual al cuadrado medio del error original, el nuevo modelo tendrá un cuadrado medio del error mayor que el original, debido a la pérdida de un grado de libertad del error. Por tanto, el nuevo modelo será en realidad peor que el original.

Hay varias concepciones erróneos en cuanto a R^2. En general, R^2 no mide la magnitud de la pendiente de la recta de regresión. Un valor grande de R^2 no implica una pendiente pronunciada. Por otra parte, R^2 no mide la adecuación del modelo, ya que puede inflarse artificialmente agregando al modelo términos polinómicos en x de orden superior. Incluso si y y x están relacionadas en una forma no lineal, R^2 con frecuencia será grande. Por ejemplo, R^2 de la ecuación de regresión de la figura 10-6*b* será relativamente grande, aun cuando la aproximación lineal es pobre. Por último, aun cuando R^2 sea grande, esto no implica necesariamente que el modelo de regresión producirá predicciones precisas de observaciones futuras.

10-8.3 Prueba de falta de ajuste

Con frecuencia los modelos de regresión se ajustan a los datos para proporcionar un modelo empírico cuando la verdadera relación entre las variables Y y x es desconocida. Naturalmente, a

uno le interesaría saber si el orden del modelo propuesto de manera tentativa es el correcto. En esta sección se describe una prueba de la validez de este supuesto.

El peligro de usar un modelo de regresión que es una aproximación pobre de la verdadera relación funcional se ilustra en la figura 10-13. Evidentemente, en esta situación debería usarse un polinomio en x de grado dos o mayor.

Se presenta ahora una prueba de la "bondad del ajuste" del modelo de regresión. Específicamente, las hipótesis que quieren probarse son

H_0: El modelo de regresión lineal simple es correcto.

H_1: El modelo de regresión lineal simple no es correcto.

La prueba implica hacer la partición de la suma de cuadrados de los errores o los residuales en los siguientes componentes:

$$SS_E = SS_{PE} + SS_{LOF} \tag{10-37}$$

donde SS_{PE} es la suma de los cuadrados atribuible al **error puro**, y SS_{LOF} es la suma de los cuadrados atribuible a la **falta de ajuste** del modelo. Para calcular SS_{PE}, deben tenerse observaciones repetidas de la respuesta Y para al menos un nivel de x. Suponga que se tienen n observaciones totales tales que

$$\begin{array}{ll} y_{11}, y_{12}, \ldots, y_{1n_1} & \text{observaciones repetidas en } x_1 \\ y_{21}, y_{22}, \ldots, y_{2n_2} & \text{observaciones repetidas en } x_2 \\ \vdots & \\ y_{m1}, y_{m2}, \ldots, y_{mn_m} & \text{observaciones repetidas en } x_m \end{array}$$

Obsérvese que hay m niveles distintos de x. La contribución a la suma de cuadrados del error puro en x_1 (por ejemplo) sería

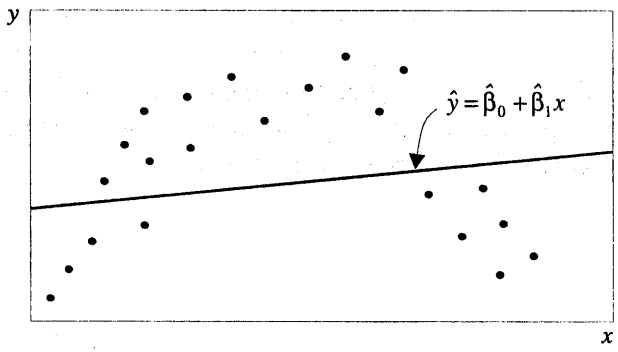


Figura 10-13 Modelo de regresión con falta de ajuste.

$$\sum_{u=1}^{n_1} (y_{1u} - \bar{y}_1)^2 \tag{10-38}$$

donde $\bar{y}_1$ representa el promedio de todas las n_1 observaciones repetidas de la respuesta y en x_1. La suma total de los cuadrados del error puro se obtendría haciendo la suma de la ecuación 10-38 en todos los niveles de x como

$$SS_{PE} = \sum_{i=1}^{m} \sum_{u=1}^{n_i} (y_{iu} - \bar{y}_i)^2 \tag{10-39}$$

Hay $n_{pe} = \sum_{i=1}^{m} (n_i - 1) = n - m$ grados de libertad asociados con la suma de los cuadrados del error puro. La suma de los cuadrados de la falta de ajuste es simplemente

$$SS_{LOF} = SS_E - SS_{PE} \tag{10-40}$$

con $n - 2 - n_{pe} = m - 2$ grados de libertad. El estadístico de la prueba de la falta de ajuste sería entonces

$$F_0 = \frac{SS_{LOF}/(m-2)}{SS_{PE}/(n-m)} = \frac{MS_{LOF}}{MS_{PE}} \tag{10-41}$$

y la hipótesis de que el modelo se ajusta adecuadamente a los datos se rechazaría si $f_0 > f_{\alpha, m-2, n-m}$.

Este procedimiento de prueba puede introducirse sin dificultad en el análisis de varianza realizado para la significación de la regresión. Si la hipótesis nula de la adecuación del modelo se rechaza, entonces el modelo debe abandonarse y deberá intentarse encontrar un modelo más apropiado. Si H_0 no es rechazada, entonces no hay razón aparente para dudar de la adecuación del modelo, y MS_{PE} y MS_{LOF} suelen combinarse para estimar σ^2.

EJEMPLO 10-9

Considérense los datos para las dos variables y y x que se presentan abajo. Ajuste un modelo de regresión lineal simple y pruebe la falta de ajuste usando $\alpha = 0.05$.

x	y
1.0	2.3, 1.8
2.0	2.8
3.3	1.8, 3.7
4.0	2.6, 2.6, 2.2
5.0	2.0

x	y
5.6	3.5, 2.8, 2.1
6.0	3.4, 3.2
6.5	3.4
6.9	5.0

Tabla 10-6 Análisis de varianza del ejemplo 10-9

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor de P
Regresión	3.4930	1	3.4930	6.66	0.0218
Residual	7.3372	14	0.5241		
(Falta de ajuste)	4.3005	7	0.6144	1.42	0.3276
(Error puro)	3.0366	7	0.4338		
Total	10.8300	15			

El modelo de regresión es $\hat{y} = 1.697 + 0.259x$, y la suma de los cuadrados de regresión es $SS_R = 3.4930$. La suma de los cuadrados del error puro se calcula como sigue:

Nivel de x	$\sum_{u=1}^{n_i}(y_{iu} - \bar{y}_i)^2$	Grados de libertad
1.0	0.1250	1
3.3	1.8050	1
4.0	0.0166	2
5.6	0.9800	2
6.0	0.0200	1
Totales	3.0366	7

El análisis de varianza se resume en la tabla 10-6. Puesto que el estadístico F de falta de ajuste es $f_0 = 1.42$, que tiene un valor P de $P = 0.3276$, no puede rechazarse la hipótesis de que el modelo tentativo describe adecuadamente los datos. Los cuadrados medios de la falta de ajuste y del error puro se agruparán para formar el cuadrado medio de los residuales, que es el cuadrado medio del denominador en la prueba de significación de la regresión. Además, puesto que el valor P para el estadístico $f_0 = 6.66$ con 1 y 14 grados de libertad, asociado con la significación de la regresión, es $P = 0.0216$, se concluye que $\beta_1 \neq 0$.

Al ajustar un modelo de regresión a datos experimentales, una buena práctica es usar el modelo con el grado menor que describa adecuadamente los datos. La prueba de falta de ajuste puede ser de utilidad a este respecto. Sin embargo, siempre es posible ajustar un polinomio de grado $n - 1$ a n puntos de datos, y el experimentador no deberá considerar el uso de un modelo que esté "saturado"; es decir, que tenga casi tantas variables independientes como observaciones de y.

EJERCICIOS DE LA SECCIÓN 10-8

10-42. Remítase a los datos del desempeño de los equipos de la NFL del ejercicio 10-4.

a) Calcule R^2 para este modelo y proponga una interpretación práctica de esta cantidad.

b) Haga una gráfica de probabilidad nornal de los residuales del modelo de mínimos cuadrados. ¿Parece satisfacerse el supuesto de normalidad?

c) Grafique los residuales contra $\hat{y}$ y contra x. Interprete estas gráficas.

10-43. Remítase a los datos del ejercicio 10-5 para el precio de venta de una casa y y los impuestos pagados x.

a) Encuentre los residuales del modelo de mínimos cuadrados.

b) Haga una gráfica de probabilidad normal de los residuales e interprétela.

c) Grafique los residuales contra $\hat{y}$ y contra x. ¿Parece satisfacerse el supuesto de que la varianza es constante?

d) ¿Qué proporción de la variabilidad total es explicada por el modelo de regresión?

10-44. En el ejercicio 10-6 se presentan los datos para y = consumo de vapor y x = temperatura mensual promedio.

a) ¿Qué proporción de la variabilidad total es explicada por el modelo de regresión lineal simple?

b) Haga una gráfica de probabilidad normal de los residuales e interpretarla.

c) Grafique los residuales contra $\hat{y}$ y contra x. ¿Parecen satisfacerse los supuestos de regresión?

10-45. Remítase a los datos del rendimiento de combustible del ejercicio 10-7.

a) ¿Qué proporción de la variabilidad total de las millas por galón de gasolina en carretera es explicada por el cilindraje del motor?

b) Grafique los residuales contra $\hat{y}$ y x, y comentar las gráficas obtenidas.

c) Haga una gráfica de probabilidad normal de los residuales. ¿Parece satisfacerse el supuesto de normalidad?

10-46. Considérense los datos del ejercicio 10-8 para y = concentración del licor verde Na_2S y x = producción de papel. Suponga que se agrega el punto muestral 14 a los datos originales, donde $y_{14} = 59$ y $x_{14} = 855$.

a) Haga un diagrama de dispersión de y contra x. Ajuste el modelo de regresión lineal simple a las 14 observaciones.

b) Pruebe la significación de la regresión con $\alpha = 0.05$.

c) Estime σ^2 para este modelo.

d) Compare la estimación de σ^2 obtenida en el inciso c) anterior con la estimación de σ^2 obtenida con los 13 puntos originales. ¿Cuál de las estimaciones es mayor y por qué?

e) Calcule los residuales para este modelo. ¿El valor de e_{14} parece extraño?

f) Haga e interprete una gráfica de probabilidad normal de los residuales.

g) Grafique los residuales contra $\hat{y}$ y contra x. Comente las gráficas obtenidas.

10-47. Remítase al ejercicio 10-9, donde se presentan los datos del aumento de la presión sanguínea y y del nivel de presión del sonido x.

a) ¿Qué proporción de la variabilidad total del aumento de la presión sanguínea es explicada por el nivel de presión del sonido?

b) Pruebe la falta de ajuste usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

c) Haga una gráfica de probabilidad normal de los residuales de este modelo de mínimos cuadrados. Interprete esta gráfica.

d) Grafique los residuales contra $\hat{y}$ y contra x. Comente las gráficas obtenidas.

10-48. En el ejercicio 10-10 se presentan los datos sobre el volumen del desgaste y y la viscosidad del aceite x.

a) Calcule R^2 para este modelo. Proponga una interpretación de esta cantidad.

b) Grafique los residuales de este modelo contra $\hat{y}$ y contra x. Interprete estas gráficas.

c) Haga una gráfica de probabilidad normal de los residuales. ¿Parece satisfacerse el supuesto de normalidad?

10-49. Remítase al ejercicio 10-11 donde se presentan los datos para la concentración de cloruros y y el área de carretera x.

a) ¿Qué proporción de la variabilidad total de la concentración de cloruros es explicada por el modelo de regresión?

b) Use las observaciones repetidas en $x = 70$ y $x = 78$ para obtener una estimación del error puro con 2 grados de libertad.

c) Use la cantidad del error puro calculada en el inciso b) para probar la falta de ajuste del modelo de regresión. Use $\alpha = 0.05$. ¿A qué conclusión puede llegarse en cuanto a la adecuación del modelo?

d) Grafique los residuales contra $\hat{y}$ y contra x. Interprete estas gráficas.

e) Haga una gráfica de probabilidad normal de los residuales. ¿Parece satisfacerse el supuesto de normalidad?

10-50. Considérense los datos de la carga propulsora de la turbina del ejercicio 10-12.

a) Calcule R^2 para este modelo. Proponga una interpretación de esta cantidad.

b) Grafique los residuales en una escala de probabilidad normal. ¿Algunos de los puntos de esta gráfica parecen extraños?

c) Elimine de la muestra los dos puntos identificados en el inciso b) y ajuste el modelo de regresión lineal simple a los 18 puntos restantes. Calcule el valor de R^2 para el nuevo modelo. ¿Es mayor o menor que el valor de R^2 calculado en el inciso a)? ¿Por qué?

d) ¿El valor de $\hat{\sigma}^2$ sufrió un cambio radical cuando se eliminaron los dos puntos identificados arriba y el modelo se ajustó a los puntos restantes? ¿Por qué?

10-51. Demuestre que una manera equivalente de definir la prueba de significación de la regresión en una regresión lineal simple es tomando R^2 como base de la prueba como sigue: para probar H_0: $\beta_1 = 0$ contra H_1: $\beta_1 \neq 0$, se calcula

$$F_0 = \frac{R^2(n-2)}{1-R^2}$$

y H_0: $\beta_1 = 0$ se rechaza si el valor calculado $f_0 > f_{\alpha, 1, n-2}$.

10-52. Suponga que el modelo de regresión lineal simple se ha ajustado a $n = 25$ observaciones y que $R^2 = 0.90$.

a) Pruebe la significación de la regresión con $\alpha = 0.05$. Use los resultados del ejercicio 10-51.

b) ¿Cuál es el valor menor de R^2 que llevaría a la conclusión de una regresión significativa si $\alpha = 0.05$?

10-53. Considérense los datos de la carga propulsora de la turbina del ejercicio 10-12. Calcule los residuales estandarizados para estos datos. ¿Esto proporciona alguna información útil acerca de la magnitud de los residuales?

10-54. **Residual studentizado**. Demuestre que la varianza del residual i-ésimo es

$$V(e_i) = \sigma^2\left[1-\left(\frac{1}{n}+\frac{(x_i-\bar{x})^2}{S_{xx}}\right)\right]$$

Sugerencia:

$$\operatorname{cov}(Y_i, \hat{Y}_i) = \sigma^2\left[\frac{1}{n}+\frac{(x_i-\bar{x})^2}{S_{xx}}\right].$$

El residual studentizado i-ésimo se define como

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2\left[1-\left(\frac{1}{n}+\frac{(x_i-\bar{x})^2}{S_{xx}}\right)\right]}}$$

a) Explique por qué r_i tiene una desviación estándar unitaria.

b) ¿Los residuales estandarizados tienen una desviación estándar unitaria?

c) Comente el comportamiento del residual estandarizado cuando el valor muestral x_i está muy cerca de la parte media del rango de x.

d) Comente el comportamiento del residual estandarizado cuando el valor muestral x_i está muy cerca de uno de los extremos del rango de x.

10-9 TRANSFORMACIONES A UNA LÍNEA RECTA

En ocasiones se encuentra que el modelo de regresión lineal $Y = \beta_0 + \beta_1 x + \varepsilon$ no es apropiado debido a que la verdadera función de regresión es no lineal. Algunas veces la naturaleza no lineal del modelo se determina por inspección visual del diagrama de dispersión, y algunas veces se sabe de antemano, como resultado de experiencias previas o de la teoría implícita, que el modelo es no lineal. En ocasiones, un diagrama de dispersión mostrará una aparente relación no lineal entre Y y x. En algunas de estas situaciones, una función no lineal puede expresarse como una línea recta mediante el uso de la transformación pertinente. A tales modelos no lineales se les llama **intrínsecamente lineales**.

Como ejemplo de un modelo no lineal que es intrínsecamente lineal, considérese la función exponencial

$$Y = \beta_0 e^{\beta_1 x} \varepsilon$$

Esta función es intrínsecamente lineal, ya que puede transformarse en una línea recta mediante una transformación logarítmica

$$ln\ Y = ln\ \beta_0 + \beta_1 x + ln\ \varepsilon$$

Esta transformación requiere que los términos transformados del error $ln\ \varepsilon$, tengan una distribución normal e independiente con media 0 y varianza σ^2.

Otra función intrínsecamente lineal es

$$Y = \beta_0 + \beta_1 \left(\frac{1}{x}\right) + \varepsilon$$

Al usar la transformación del recíproco $z = 1/x$, el modelo se hace lineal, y se obtiene

$$Y = \beta_0 + \beta_1 z + \varepsilon$$

En ocasiones pueden emplearse varias transformaciones de manera conjunta para hacer lineal una función. Por ejemplo, considérese la función

$$Y = \frac{1}{\exp(\beta_0 + \beta_1 x + \varepsilon)}$$

Al hacer $Y^* = 1/Y$, se llega a la forma lineal

$$ln\ Y^* = \beta_0 + \beta_1 x + \varepsilon$$

Daniel y Wood ofrecen varios modelos no lineales que son intrínsecamente lineales. Para ejemplos del ajuste de dichos modelos, consulte Montgomery y Peck, o Myers.

10-10 CORRELACIÓN

En el análisis de regresión desarrollado aquí se ha supuesto que x es una variable matemática, medida con error omisible, y que Y es una variable aleatoria. Muchas aplicaciones del análisis de regresión incluyen situaciones en las que tanto X como Y son variables aleatorias. En estas situaciones, es común suponer que las observaciones (X_i, Y_i), $i = 1, 2, ..., n$, son variables aleatorias con una distribución conjunta que se obtiene de la distribución $f(x, y)$.

Por ejemplo, suponga que quiere desarrollarse un modelo de regresión que relacione la resistencia al corte de los puntos de soldadura con el diámetro de los mismos. En este ejemplo, el diámetro de la soldadura no puede controlarse. Al analista le gustaría seleccionar aleatoriamente n puntos de soldadura y observar el diámetro (X_i) y la resistencia al corte (Y_i) de cada uno de ellos. Por lo tanto, (X_i, Y_i) son variables aleatorias con una distribución conjunta.

Se supone que la distribución conjunta de X_i y Y_i es la distribución normal de dos variables presentada en el capítulo 6, y que μ_Y y σ_Y^2 son la media y la varianza de Y, que μ_X y σ_X^2 son la media y la varianza de X, y que ρ es el coeficiente de correlación entre Y y X. Recuérdese que el coeficiente de correlación se define como

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{10-42}$$

donde σ_{XY} es la covarianza entre Y y X.

La distribución condicional de Y para un valor dado de $X = x$ es

$$f_{Y|x}(y) = \frac{1}{\sqrt{2\pi}\sigma_{Y|x}} \exp\left[-\frac{1}{2}\left(\frac{y - \beta_0 + \beta_1 x}{\sigma_{Y|x}}\right)^2\right] \tag{10-43}$$

donde

$$\beta_0 = \mu_Y - \mu_X \rho \frac{\sigma_Y}{\sigma_X} \tag{10-44}$$

$$\beta_1 = \frac{\sigma_Y}{\sigma_X} \rho \tag{10-45}$$

y la varianza de la distribución condicional de Y dada $X = x$ es

$$\sigma_{Y|x}^2 = \sigma_Y^2 (1 - \rho^2) \tag{10-46}$$

Es decir, la distribución condicional de Y dada $X = x$ es normal con media

$$E(Y|x) = \beta_0 + \beta_1 x \tag{10-47}$$

y varianza $\sigma_{Y|x}^2$. Por tanto, la media de la distribución condicional de Y dada $X = x$ es un modelo de regresión lineal simple. Además, existe una relación entre el coeficiente de correlación ρ y la

pendiente β_1. Por la ecuación 10-45 se observa que si $\rho = 0$, entonces $\beta_1 = 0$, lo cual implica que no hay regresión de Y sobre X. Es decir, el conocimiento de X no ayuda a predecir Y.

Puede usarse el método de máxima verosimilitud para estimar los parámetros β_0 y β_1. Es posible demostrar que los estimadores de máxima verosimilitud de dichos parámetros son

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \tag{10-48}$$

y

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i (X_i - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{S_{XY}}{S_{XX}} \tag{10-49}$$

Se observa que los estimadores de la ordenada al origen y la pendiente de las ecuaciones 10-48 y 10-49 son idénticos a los que se obtuvieron con el método de mínimos cuadrados en el caso en que se supuso que X era una variable matemática. Es decir, el modelo de regresión cuando Y y X tienen una distribución normal conjunta es equivalente al modelo cuando X se considera una variable matemática. Esto se debe a que las variables aleatorias Y dada $X = x$ tienen una distribución normal e independiente con media $\beta_0 + \beta_1 x$ y varianza constante $\sigma^2_{Y|x}$. Estos resultados seguirán siendo válidos para cualquier distribución conjunta de Y y X tal que la distribución condicional de Y dada X es normal.

Es posible hacer inferencias acerca del coeficiente de correlación ρ de este modelo. El estimador ρ es el **coeficiente de correlación muestral**

$$R = \frac{\sum_{i=1}^{n} Y_i (X_i - \bar{X})}{\left[\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y_i - \bar{Y})^2\right]^{1/2}} = \frac{S_{XY}}{[S_{XX} S_{YY}]^{1/2}} \tag{10-50}$$

Obsérvese que

$$\hat{\beta}_1 = \left(\frac{S_{XY}}{S_{XX}}\right)^{1/2} R \tag{10-51}$$

por lo que la pendiente $\hat{\beta}_1$ es el coeficiente de correlación muestral R multiplicado por un factor escalar que es la raíz cuadrada de "lo esparcido" de los valores de Y dividida por "lo esparcido" de los valores de X. Por tanto, $\hat{\beta}_1$ y R guardan una estrecha relación, aun cuando proporcionan información un tanto diferente. El coeficiente de correlación muestral R mide la asociación lineal entre Y y X, mientras que $\hat{\beta}_1$ mide el cambio predicho en la media de Y para un cambio

unitario en X. En el caso de una variable matemática x, R no tiene significado debido a que la magnitud de R depende de la elección del espaciamiento de x. También, por la ecuación 10-51, puede escribirse

$$\begin{aligned} R^2 &= \hat{\beta}_1^2 \frac{S_{XX}}{S_{YY}} \\ &= \frac{\hat{\beta}_1 S_{XY}}{S_{YY}} \\ &= \frac{SS_R}{S_{YY}} \end{aligned}$$

que se identifica, por la ecuación 10-36, como el coeficiente de determinación. Es decir, el coeficiente de determinación R^2 no es sino el cuadrado del coeficiente de correlación entre Y y X.

Con frecuencia es útil probar las hipótesis

$$\begin{aligned} H_0&: \rho = 0 \\ H_1&: \rho \neq 0 \end{aligned} \tag{10-52}$$

El estadístico de la prueba apropiado para estas hipótesis es

$$T_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} \tag{10-53}$$

que tiene la distribución t con $n-2$ grados de libertad si H_0: $\rho = 0$ es verdadera. Por lo tanto, la hipótesis nula se rechazaría si $|t_0| > t_{\alpha/2,\, n-2}$. Esta prueba es equivalente a la prueba de la hipótesis H_0: $\beta_1 = 0$ dada en la sección 10-5.1. Esta equivalencia se sigue directamente de la ecuación 10-51.

El procedimiento para probar la hipótesis

$$\begin{aligned} H_0&: \rho = \rho_0 \\ H_1&: \rho \neq \rho_0 \end{aligned} \tag{10-54}$$

donde $\rho_0 \neq 0$ es un tanto más complicada. Para muestras moderadamente grandes (por ejemplo, $n \geq 25$), el estadístico

$$Z = \text{arctanh } R = \frac{1}{2} \ln \frac{1+R}{1-R} \tag{10-55}$$

tiene una distribución aproximadamente normal con media

$$\mu_Z = \text{arctanh } \rho = \frac{1}{2} ln \frac{1+\rho}{1-\rho}$$

y varianza

$$\sigma_Z^2 = (n-3)^{-1}$$

Por lo tanto, para probar la hipótesis H_0: $\rho = \rho_0$, puede usarse el estadístico de la prueba

$$Z_0 = (\text{arctanh } R - \text{arctanh } \rho_0)(n-3)^{1/2} \qquad (10\text{-}56)$$

y H_0: $\rho = \rho_0$ se rechaza si el valor del estadístico de la prueba de la ecuación 10-56 es tal que $|z_0| > z_{\alpha/2}$.

También es posible construir un intervalo de confianza del $100(1 - \alpha)$ por ciento aproximado para ρ utilizando la transformación de la ecuación 10-55. El intervalo de confianza del $100(1 - \alpha)$ por ciento aproximado es

$$\tanh\left(\text{arctanh } r - \frac{z_{\alpha/2}}{\sqrt{n-3}}\right) \le \rho \le \tanh\left(\text{arctanh } r + \frac{z_{\alpha/2}}{\sqrt{n-3}}\right) \qquad (10\text{-}57)$$

donde $\tanh u = (e^u - e^{-u})/(e^u + e^{-u})$.

EJEMPLO 10-10

En el capítulo 1 (sección 1-4) se describe una aplicación del análisis de regresión en la que un ingeniero en una planta de ensamblaje de semiconductores investiga la relación entre la resistencia al desprendimiento de un alambre adherido y dos factores: la longitud del alambre y la altura del broquel. En este ejemplo, sólo se considerará uno de los factores, la longitud del alambre. Se selecciona y se prueba una muestra aleatoria de 25 unidades, observándose la resistencia al desprendimiento del alambre adherido y la longitud del alambre en cada unidad. Los datos se presentan en la tabla 1-2, página 15. Se supone que la resistencia al desprendimiento y la longitud del alambre siguen una distribución normal conjunta.

Usando los datos de la tabla 1-2, pueden calcularse

$$SS_T = 6\ 105.9447 \quad S_{xx} = 698.5600 \quad \text{y} \quad S_{xy} = 2\ 027.7132$$

El modelo de regresión es

$$\hat{y} = 5.1145 + 2.9027x$$

El coeficiente de correlación muestral entre X y Y se calcula con la ecuación 10-50 como

$$r = \frac{S_{xy}}{[S_{xx}S_{yy}]^{1/2}} = \frac{2\,027.7132}{[(698.560)(6\,105.9447)]^{1/2}} = 0.9818$$

Obsérvese que $r^2 = (0.9818)^2 = 0.9640$, o que aproximadamente 96.40% de la variabilidad de la resistencia al desprendimiento está explicada por la relación lineal con la longitud del alambre.

Suponga ahora que quiere probarse la hipótesis

$$H_0\colon \rho = 0$$
$$H_1\colon \rho \neq 0$$

con $\alpha = 0.05$. El estadístico t de la ecuación 10-53 puede calcularse como

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.9818\sqrt{23}}{\sqrt{1-0.9640}} = 24.82$$

y puesto que $t_{0.025,\,23} = 2.069$, se rechaza H_0 y se concluye que el coeficiente de correlación $\rho \neq 0$.

Por último, a partir de la ecuación 10-57 puede construirse un intervalo de confianza de 95% aproximado para ρ. Puesto que arctanh r = arctanh 0.9818 = 2.3452, la ecuación 10-57 es

$$\tanh\left(2.3452 - \frac{1.96}{\sqrt{22}}\right) \leq \rho \leq \tanh\left(2.3452 + \frac{1.96}{\sqrt{22}}\right)$$

que se reduce a

$$0.9585 \leq \rho \leq 0.9921$$

EJERCICIOS DE LA SECCIÓN 10-10

10-55. Se presentan a continuación el promedio del examen final y los exámenes parciales de 20 estudiantes seleccionados al azar que llevaron un curso de estadística para ingenieros y un curso de investigación de operaciones (IO). Suponga que los promedios finales tienen una distribución normal conjunta.

a) Encuentre la recta de regresión que relacione el promedio final de estadística con el promedio final de IO.

b) Pruebe la significación de la regresión usando $\alpha = 0.05$.

Estadística	86	75	69	75	90
IO	80	81	75	81	92
Estadística	94	83	86	71	65
IO	95	80	81	76	72
Estadística	84	71	62	90	83
IO	85	72	65	93	81
Estadística	75	71	76	84	97
IO	70	73	72	80	98

c) Estime el coeficiente de correlación.
d) Pruebe la hipótesis de que $\rho = 0$ usando $\alpha = 0.05$.
e) Pruebe la hipótesis de que $\rho = 0.5$ usando $\alpha = 0.05$.
f) Construya un intervalo de confianza de 95% para el coeficiente de correlación.

10-56. En la tabla siguiente se muestra el peso y la presión sanguínea sistólica de 26 hombres seleccionados al azar del grupo de edad entre 25 y 30 años. Suponga que el peso y la presión sanguínea tienen una distribución normal conjunta.

Sujeto	Peso	Presión sanguínea sistólica
1	165	130
2	167	133
3	180	150
4	155	128
5	212	151
6	175	146
7	190	150
8	210	140
9	200	148
10	149	125
11	158	133
12	169	135
13	170	150
14	172	153
15	159	128
16	168	132
17	174	149
18	183	158
19	215	150
20	195	163
21	180	156
22	143	124
23	240	170
24	235	165
25	192	160
26	187	159

a) Encuentre una recta de regresión que relacione la presión sanguínea sistólica con el peso.
b) Pruebe la significación de la regresión usando $\alpha = 0.05$.
c) Estime el coeficiente de correlación.
d) Pruebe la hipótesis de que $\rho = 0$ usando $\alpha = 0.05$.
e) Pruebe la hipótesis de que $\rho = 0.6$ usando $\alpha = 0.05$.
f) Construya un intervalo de confianza de 95% para el coeficiente de correlación.

10-57. Considérense los datos de la NFL introducidos en el ejercicio 10-4.
a) Estime el coeficiente de correlación entre el número de juegos ganados y las yardas corridas en el campo de los oponentes.
b) Pruebe la hipótesis H_0: $\rho = 0$ contra H_1: $\rho \neq 0$ usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?
c) Construya un intervalo de confianza de 95% para ρ.
d) Pruebe la hipótesis H_0: $\rho = -0.7$ contra H_1: $\rho \neq -0.7$ usando $\alpha = 0.05$. Encuentre el valor P para esta prueba.

10-58. Demuestre que el estadístico t de la ecuación 10-53 para probar H_0: $\rho = 0$ es idéntico al estadístico t para probar H_0: $\beta_1 = 0$.

10-59. Se tomó una muestra aleatoria de 50 observaciones del diámetro de los puntos de soldadura y de la resistencia al corte de la soldadura correspondiente.
a) Dado que $r = 0.62$, pruebe la hipótesis de que $\rho = 0$ usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?
b) Encuentre un intervalo de confianza de 99% para ρ.
c) Con base en el intervalo de confianza del inciso b), ¿puede concluirse que $\rho = 0.5$ con el nivel de significación 0.01?

10-60. Suponga que una muestra aleatoria de 10 000 pares (X, Y) dio como resultado un coeficiente de correlación muestral de $r = 0.02$.
a) ¿A qué conclusión se llegaría si se probara H_0: $\rho = 0$ usando $\alpha = 0.05$? ¿Cuál es el valor P para esta prueba?
b) Comente la significación práctica y la significación estadística de la respuesta del inciso a).

10-61. Los datos que se presentan abajo corresponden a X = el contenido de agua de nieve el 1° de abril y Y = el aforo de abril a julio (en pulgadas) en la vertiente del río Snake en Wyoming de 1919 a 1935. (Los datos se tomaron de un artículo de *Research Notes,* Vol. 61. Pacific Northwest Forest Range Experiment Station, Oregon.)

x	y	x	y
23.1	10.5	37.9	22.8
32.8	16.7	30.5	14.1
31.8	18.2	25.1	12.9
32.0	17.0	12.4	8.8
30.4	16.3	35.1	17.4
24.0	10.5	31.5	14.9
39.5	23.1	21.1	10.5
24.2	12.4	27.6	16.1
52.5	24.9		

a) Estime la correlación entre Y y X.
b) Pruebe la hipótesis de que $\rho = 0$ usando $\alpha = 0.05$.
c) Ajuste un modelo de regresión lineal simple y pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse? ¿En qué forma se relaciona la prueba de significación de la regresión con la prueba de ρ del inciso b) anterior?
d) Pruebe la hipótesis H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ y saque conclusiones. Use $\alpha = 0.05$.
e) Analice los residuales y comente la adecuación del modelo.

10-62. Se tomó una muestra aleatoria de $n = 25$ observaciones del tiempo hasta la falla de un componente electrónico y de la temperatura en el ambiente de aplicación en que se usaba el componente.
a) Dado que $r = 0.83$, pruebe la hipótesis de que $\rho = 0$ usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?
b) Encuentre un intervalo de confianza de 95% para ρ.
c) Pruebe la hipótesis H_0: $\rho = 0.8$ contra H_1: $\rho \neq 0.8$ usando $\alpha = 0.05$. Encuentre el valor P para esta prueba.

Ejercicios complementarios

10-63. Demuestre que, para el modelo de regresión lineal simple, los siguientes enunciados son verdaderos:

a) $\sum_{i=1}^{n} (y_i - \hat{y}_i) = 0$

b) $\sum_{i=1}^{n} (y_i - \hat{y}_i)x_i = 0$

c) $\frac{1}{n}\sum_{i=1}^{n} \hat{y}_i = y$

10-64. En un artículo de *IEEE Transactions on Instrumentation and Measurement* ("Medición directa, rápida y precisa de V_T y K del transistor MOS usando un circuito V_T-Sift", Vol. 40, pp. 951-955) se describe el uso de un modelo de regresión lineal simple para expresar la corriente consumida y (en mA) como una función del voltaje tierra a fuente x (en V). Los datos son los siguientes:

y	x
0.734	1.1
0.886	1.2
1.04	1.3
1.19	1.4
1.35	1.5
1.50	1.6
1.66	1.7
1.81	1.8
1.97	1.9
2.12	2.0

a) Trace un diagrama de dispersión de estos datos. ¿Parece plausible una relación lineal?
b) Ajuste un modelo de regresión lineal simple a estos datos.

c) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

d) Encuentre un intervalo de confianza de 95% para la pendiente.

e) Probar la hipótesis H_0: $\beta_0 = 0$ contra H_1: $\beta_0 \neq 0$ usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?

10-65. La resistencia del papel usado para hacer cajas de cartón (y) está relacionada con el porcentaje de la concentración de madera dura en la pulpa original (x). Bajo condiciones controladas, se fabrican 16 muestras en una planta piloto, usando en cada una un lote de pulpa diferente, y se mide la resistencia de tensión. Los datos se presentan a continuación:

y	101.4	117.4	117.1	106.2
x	1.0	1.5	1.5	1.5
y	131.9	146.9	146.8	133.9
x	2.0	2.0	2.2	2.4
y	111.0	123.0	125.1	145.2
x	2.5	2.5	2.8	2.8
y	134.3	144.5	143.7	146.9
x	3.0	3.0	3.2	3.3

a) Ajuste el modelo de regresión lineal simple a estos datos.

b) Pruebe la significación de la regresión usando $\alpha = 0.05$.

c) Construya un intervalo de confianza de 90% para la pendiente β_1.

d) Construya un intervalo de confianza de 90% para la ordenada al origen β_0.

e) Construir un intervalo de confianza de 95% para la resistencia media en $x = 2.5$.

f) Pruebe la falta de ajuste usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

10-66. Los datos siguientes corresponden a la generación de corriente directa (CD) de un generador eólico (y) y la velocidad del viento (x).

Número de observación, i	Velocidad del viento (MPH), x_i	Salida de CD, y_i
1	5.00	1.582
2	6.00	1.822
3	3.40	1.057
4	2.70	0.500
5	10.00	2.236
6	9.70	2.386
7	9.55	2.294
8	3.05	0.558
9	8.15	2.166
10	6.20	1.866
11	2.90	0.653
12	6.35	1.930
13	4.60	1.562
14	5.80	1.737
15	7.40	2.088
16	3.60	1.137
17	7.85	2.179
18	8.80	2.112
19	7.00	1.800
20	5.45	1.501
21	9.10	2.303
22	10.20	2.310
23	4.10	1.194
24	3.95	1.144
25	2.45	0.123

a) Trace un diagrama de dispersión de estos datos. ¿Qué tipo de relación parece apropiada para relacionar y con x?

b) Ajuste el modelo de regresión lineal simple a estos datos.

c) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?

d) Grafique los residuales del modelo de regresión lineal simple contra $\hat{y}_i$ y contra la velocidad del viento x. ¿A qué conclusión se llega en cuanto a la adecuación del modelo?

e) Trace un diagrama de dispersión de y contra $1/x$. ¿Qué tipo de relación parece apropiado?

f) Ajuste el modelo de regresión lineal simple que relacione y con $1/x$. Pruebe la significación de la regresión (use $\alpha = 0.05$), y grafique los residuales de este modelo contra $\hat{y}_i$ y $1/x_i$. ¿Qué puede concluirse en cuanto a la adecuación del modelo?

10-67. Una empresa que genera energía eléctrica está interesada en desarrollar un modelo que relacione la demanda en las horas pico (y, en kw) con el consumo mensual total de energía durante el mes (x, en kwh). Los datos de 50 consumidores residenciales se presentan en la tabla de la columna derecha.

a) Trace un diagrama de dispersión de y contra x.

b) Ajuste el modelo de regresión lineal simple.

c) Pruebe la significación de la regresión usando $\alpha = 0.05$.

d) Grafique los residuales contra $\hat{y}_i$ y comente los supuestos de regresión implícitos. Específicamente, ¿parece satisfacerse el supuesto de la igualdad de la varianza?

e) Encuentre un modelo de regresión lineal simple usando $\sqrt{y}$ como variable de respuesta. ¿Esta transformación de y estabiliza el problema de la desigualdad de la varianza detectado en el inciso d) anterior?

10-68. Considérense los datos que se presentan abajo. Suponga que se considera que la relación entre Y y x es $Y = (\beta_0 + \beta_1 x + \varepsilon)^{-1}$. Ajuste el modelo apropiado a estos datos. ¿La forma del modelo supuesto parece razonable?

x	10	15	18	12
y	0.1	0.13	0.09	0.15

x	9	8	11	6
y	0.20	0.21	0.18	0.24

10-69. Considérense los datos del peso y la presión sanguínea del ejercicio 10-56. Ajuste un modelo sin ordenada al origen a los datos, y compárelo con el modelo obtenido en el ejercicio 10-56. ¿Cuál de los modelos es mejor?

Cliente	x(kwh)	y(kw)
1	679	0.79
2	292	0.44
3	1 012	0.56
4	493	0.79
5	582	2.70
6	1 156	3.64
7	997	4.73
8	2 189	9.50
9	1 097	5.34
10	2 078	6.85
11	1 818	5.84
12	1 700	5.21
13	747	3.25
14	2 030	4.43
15	1 643	3.16
16	414	0.50
17	354	0.17
18	1 276	1.88
19	745	0.77
20	795	3.70
21	540	0.56
22	874	1.56
23	1 543	5.28
24	1 029	0.64
25	710	4.00
26	1 434	0.31
27	837	4.20
28	1 748	4.88
29	1 381	3.48
30	1 428	7.58
31	1 255	2.63
32	1 777	4.99
33	370	0.59
34	2 316	8.19
35	1 130	4.79
36	463	0.51
37	770	1.74
38	724	4.10
39	808	3.94
40	790	0.96
41	783	3.29
42	406	0.44
43	1 242	3.24
44	658	2.14
45	1 746	5.71
46	895	4.12
47	1 114	1.90
48	413	0.51
49	1 787	8.33
50	3 560	14.94

10-70. Los datos que se presentan abajo, adaptados de Montgomery y Peck, corresponden al número de deficiencias mentales certificadas por cada 10 000 habitantes de la población estimada del Reino Unido (y) y al número de concesiones de recepción de radio (x) otorgadas por la BBC (en millones) en los años de 1924 a 1937. Haga el ajuste de un modelo de regresión lineal simple que relacione y y x. Comente el modelo. Específicamente, ¿la existencia de una fuerte correlación implica una relación de causa y efecto?

Año	Número de deficiencias mentales certificadas por cada 10 000 habitantes de la población estimada del Reino Unido (y)	Número de concesiones de recepción de radio otorgadas (en millones) en el Reino Unido (x)
1924	8	1.350
1925	8	1.960
1926	9	2.270
1927	10	2.483
1928	11	2.730
1929	11	3.091
1930	12	3.674
1931	16	4.620
1932	18	5.497
1933	19	6.260
1934	20	7.012
1935	21	7.618
1936	22	8.131
1937	23	8.593

EJERCICIOS PARA DESARROLLAR EL INTELECTO

10-71. Considérese el modelo de regresión lineal simple $Y = \beta_0 + \beta_1 x + \varepsilon$, con $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$, y los errores ε no correlacionados.

a) Demuestre que $\text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\bar{x}\,\sigma^2/S_{xx}$.

b) Demuestre que $\text{cov}(\bar{Y}, \hat{\beta}_1) = 0$.

10-72. Considérese el modelo de regresión lineal simple $Y = \beta_0 + \beta_1 x + \varepsilon$, con $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$, y los errores ε no correlacionados.

a) Demuestre que $E(\hat{\sigma}^2) = E(MS_E) = \sigma^2$.

b) Demuestre que $E(MS_R) = \sigma^2 + \beta_1^2 S_{xx}$.

10-73. Considere que se ha supuesto el modelo de regresión lineal

$$Y = \beta_0 + \beta_1 x_1 + \varepsilon$$

pero la respuesta está afectada por una segunda variable x_2 tal que la verdadera función de regresión es

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

¿El estimador de la pendiente del modelo de regresión lineal simple es insesgado?

10-74. Suponga que se está ajustando una recta y quiere hacerse la varianza del coeficiente de regresión $\hat{\beta}_1$ tan pequeña como sea posible. ¿De dónde deberán tomarse las observaciones x_i, $i = 1, 2, \ldots, n$, para minimizar $V(\hat{\beta}_1)$? Discuta las implicaciones prácticas de esta asignación de las x_i.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

10-75. Cuadrados medios ponderados. Suponga que se está ajustando la recta $Y = \beta_0 + \beta_1 x + \varepsilon$, pero la varianza de Y depende del nivel de x; es decir

$$V(Y_i|x_i) = \sigma_i^2 = \frac{\sigma^2}{w_i} \quad i = 1, 2, \ldots, n$$

donde las w_i son constantes, a las que con frecuencia se les llama *ponderaciones*. Demuestre que para una función objetivo apropiada, las ecuaciones normales de mínimos cuadrados resultantes son

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i + \hat{\beta}_1 \sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i x_i + \hat{\beta}_1 \sum_{i=1}^{n} w_i x_i^2 = \sum_{i=1}^{n} w_i x_i y_i$$

10-76. Considérese una situación en la que tanto Y como X son variables aleatorias. Sean s_x y s_y las desviaciones estándar muestrales de las x y y observadas, respectivamente. Demuestre que una expresión alternativa para el modelo de regresión lineal simple ajustado $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ es

$$\hat{y} = \bar{y} + r\frac{s_y}{s_x}(x - \bar{x})$$

10-77. Suponga que se tiene interés en ajustar un modelo de regresión lineal simple $Y = \beta_0 + \beta_1 x + \varepsilon$, donde la ordenada al origen, β_0, es conocida.

a) Encuentre el estimador de mínimos cuadrados de β_1.

b) ¿Cuál es la varianza del estimador de la pendiente del inciso a)?

c) Encuentre una expresión para un intervalo de confianza del $100(1 - \alpha)$ por ciento para la pendiente β_1. ¿Este intervalo tiene una longitud mayor que el intervalo correspondiente al caso en que son desconocidas tanto la ordenada al origen como la pendiente? Justifique su respuesta.

11 Regresión lineal múltiple

PLAN GENERAL DEL CAPÍTULO

11-1 MODELO DE REGRESIÓN LINEAL MÚLTIPLE

11-2 ESTIMACIÓN DE LOS PARÁMETROS POR MÍNIMOS CUADRADOS

11-3 ENFOQUE MATRICIAL DE LA REGRESIÓN LINEAL MÚLTIPLE

11-4 PROPIEDADES DE LOS ESTIMADORES DE MÍNIMOS CUADRADOS Y ESTIMACIÓN DE σ^2

11-5 PRUEBAS DE HIPÓTESIS EN LA REGRESIÓN LINEAL MÚLTIPLE

- 11-5.1 Prueba para la significación de una regresión
- 11-5.2 Pruebas para los coeficientes de regresión individuales y para subconjuntos de coeficientes

11-6 INTERVALOS DE CONFIANZA EN LA REGRESIÓN LINEAL MÚLTIPLE

- 11-6.1 Intervalos de confianza para los coeficientes individuales de una regresión
- 11-6.2 Intervalo de confianza para la respuesta media

11-7 PREDICCIÓN DE NUEVAS OBSERVACIONES

11-8 MEDIDAS DE LA ADECUACIÓN DEL MODELO

- 11-8.1 Coeficiente de determinación múltiple (R^2)
- 11-8.2 Análisis de residuales
- 11-8.3 Observaciones influyentes

11-9 MODELOS DE REGRESIÓN POLINOMIAL

11-10 VARIABLES INDICADORAS

11-11 SELECCIÓN DE VARIABLES EN REGRESIONES MÚLTIPLES

- 11-11.1 Problema de la construcción del modelo
- 11-11.2 Procedimientos computacionales para la selección de variables
- 11-11.3 Salida de computadora de una regresión por pasos

11-12 COLINEALIDAD MÚLTIPLE

11-1 MODELO DE REGRESIÓN LINEAL MÚLTIPLE

Muchas aplicaciones del análisis de regresión incluyen situaciones en las que hay más de una variable regresora. A un modelo de regresión que contiene más de una variable regresora se le llama **modelo de regresión múltiple**.

Como un ejemplo, suponga que la vida efectiva de una herramienta de corte depende de la velocidad de corte y del ángulo de la herramienta. Un modelo de regresión múltiple que podría describir esta relación es

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \tag{11-1}$$

donde Y representa la vida de la herramienta, x_1 representa la velocidad de corte, x_2 representa el ángulo de la herramienta, y ε es un término del error aleatorio. Se trata de un **modelo de regresión lineal múltiple** con dos regresores o variables independientes. Se usa el término *lineal* porque la ecuación 11-1 es una función lineal de los parámetros desconocidos β_0, β_1 y β_2.

El modelo de regresión de la ecuación 11-1 describe un plano en el espacio tridimensional de Y, x_1 y x_2. En la figura 11-1*a* se ilustra este plano para el modelo de regresión

$$E(Y) = 50 + 10x_1 + 7x_2$$

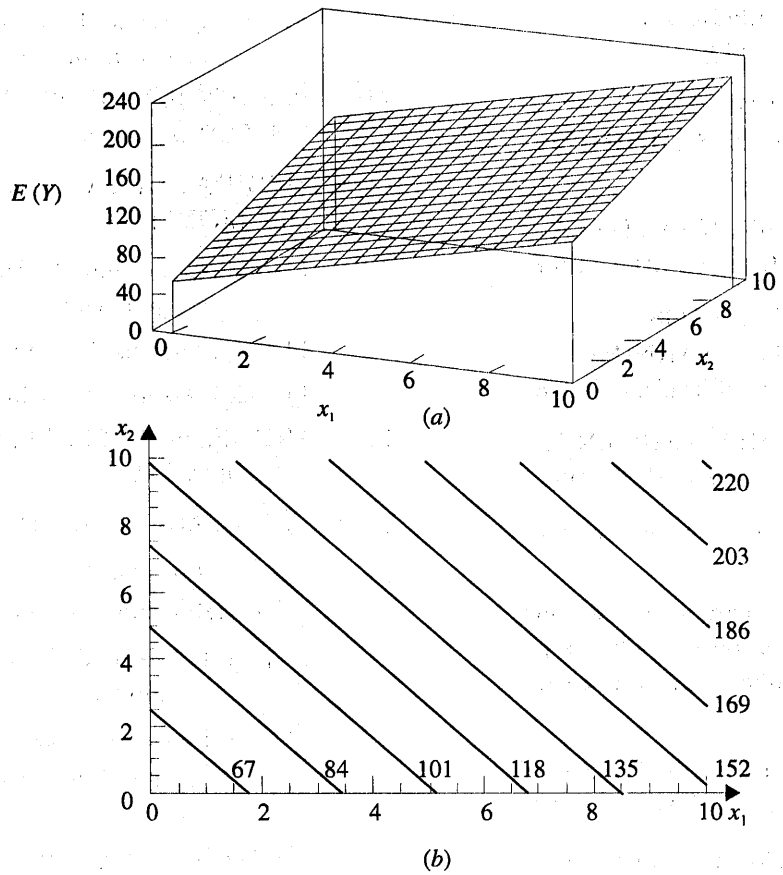


Figura 11-1 ***a***) El plano de regresión del modelo $E(Y) = 50 + 10x_1 + 7x_2$. ***b***) La gráfica de contorno.

donde se ha supuesto que el valor esperado del término del error es cero; es decir, $E(\varepsilon) = 0$. El parámetro β_0 es la **ordenada al origen** del plano. En ocasiones a β_1 y β_2 se les llama los **coeficientes de regresión parciales**, debido a que β_1 mide el cambio esperado en Y por un cambio unitario de x_1 cuando x_2 se mantiene constante, y β_2 mide el cambio esperado en Y por un cambio unitario de x_2 cuando x_1 se mantiene constante. En la figura 11-1*b* se muestra la **gráfica de contorno** (curvas de nivel) del modelo de regresión; es decir, las líneas de $E(Y)$ constante como una función de x_1 y x_2. Obsérvese que las líneas de contorno de esta gráfica son rectas.

En general, la **variable de respuesta** o **dependiente** y, puede relacionarse con k **variables regresoras** o **independientes**. Al modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{11-2}$$

se le llama modelo de regresión lineal múltiple con k variables regresoras. A los parámetros β_j, $j = 0, 1, \ldots, k$, se les llama los coeficientes de regresión. Este modelo describe un hiperplano en el espacio de k dimensiones de las variables regresoras $\{x_j\}$. El parámetro β_j representa el cambio esperado en la variable de respuesta Y, por un cambio unitario en x_j cuando se mantienen constantes las demás variables regresoras x_i $(i \neq j)$.

Los modelos de regresión lineal múltiple se usan con frecuencia como funciones de aproximación. Es decir, se desconoce la verdadera relación funcional entre Y y $x_1, x_2, \ldots, x_k$, pero en ciertos rangos de las variables independientes el modelo de regresión lineal es una aproximación adecuada.

Muchas veces, incluso los modelos cuya estructura es más compleja que la ecuación 11-2, pueden analizarse mediante técnicas de regresión lineal múltiple. Por ejemplo, considérese el modelo de un polinomio cúbico en una variable regresora.

$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon \tag{11-3}$$

Si se hace $x_1 = x$, $x_2 = x^2$, $x_3 = x^3$, entonces la ecuación 11-3 puede escribirse como

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{11-4}$$

que es un modelo de regresión lineal múltiple con tres variables regresoras.

Los modelos que incluyen efectos **de interacción** también pueden analizarse utilizando métodos de regresión lineal múltiple. Una interacción entre dos variables puede representarse con un término con un producto cruzado en el modelo, tal como

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \tag{11-5}$$

Si se hace $x_3 = x_1 x_2$ y $\beta_3 = \beta_{12}$, entonces la ecuación 11-5 puede escribirse como

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

que es un modelo de regresión lineal.

En la figura 11-2*a* y *b* se muestra la gráfica tridimensional del modelo de regresión

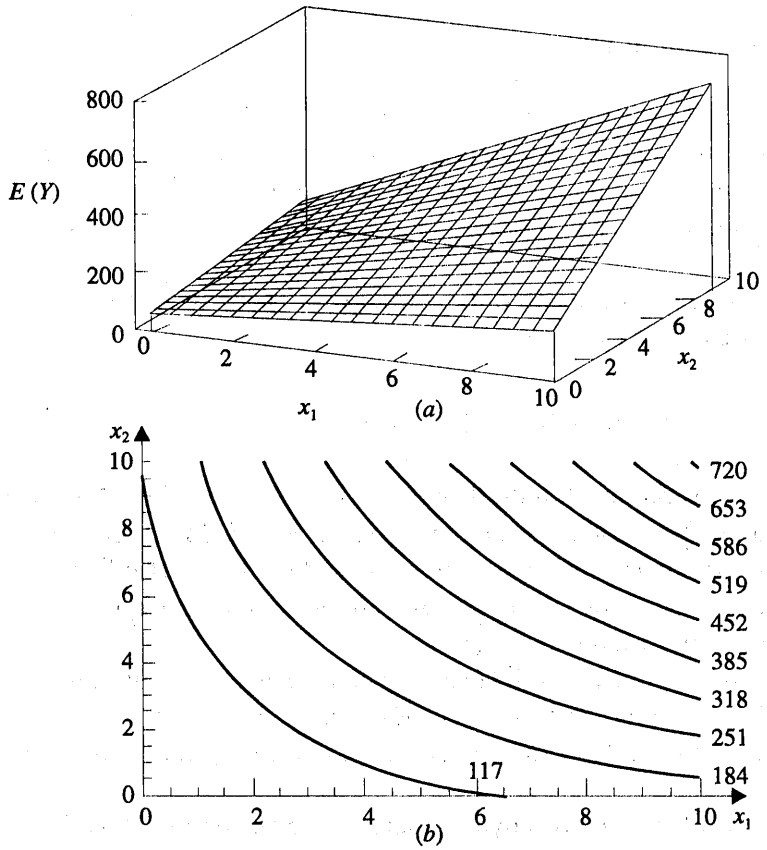


Figura 11-2 *a*) Gráfica tridimensional del modelo de regresión $E(Y) = 50 + 10x_1 + 7x_2 + 5x_1x_2$. *b*) La gráfica de contorno.

$$Y = 50 + 10x_1 + 7x_2 + 5x_1x_2$$

y la gráfica de contorno bidimensional (curvas de nivel) correspondiente. Obsérvese que, no obstante que este modelo es un modelo de regresión lineal, la forma de la superficie que genera el modelo es no lineal. En general, **cualquier modelo de regresión cuyos parámetros** (las β) **son lineales es un modelo de regresión lineal, independientemente de la forma de la superficie que genera.**

En la figura 11-2 se muestra una ilustrativa interpretación gráfica de una interacción. En general, una interacción implica que el efecto producido por el cambio en una variable (x_1, por ejemplo) depende del nivel de la otra variable (x_2). Por ejemplo, en la figura 11-2 se muestra que un cambio de x_1 de 2 a 8 produce un cambio mucho menor en $E(Y)$ cuando $x_2 = 2$ que cuando $x_2 = 10$. Los efectos de interacción son comunes en el estudio y análisis de sistemas del mundo real, y los métodos de regresión son una de las técnicas que pueden usarse para describirlos.

Como ejemplo final, considérese el modelo de segundo orden con interacción

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon \qquad (11\text{-}6)$$

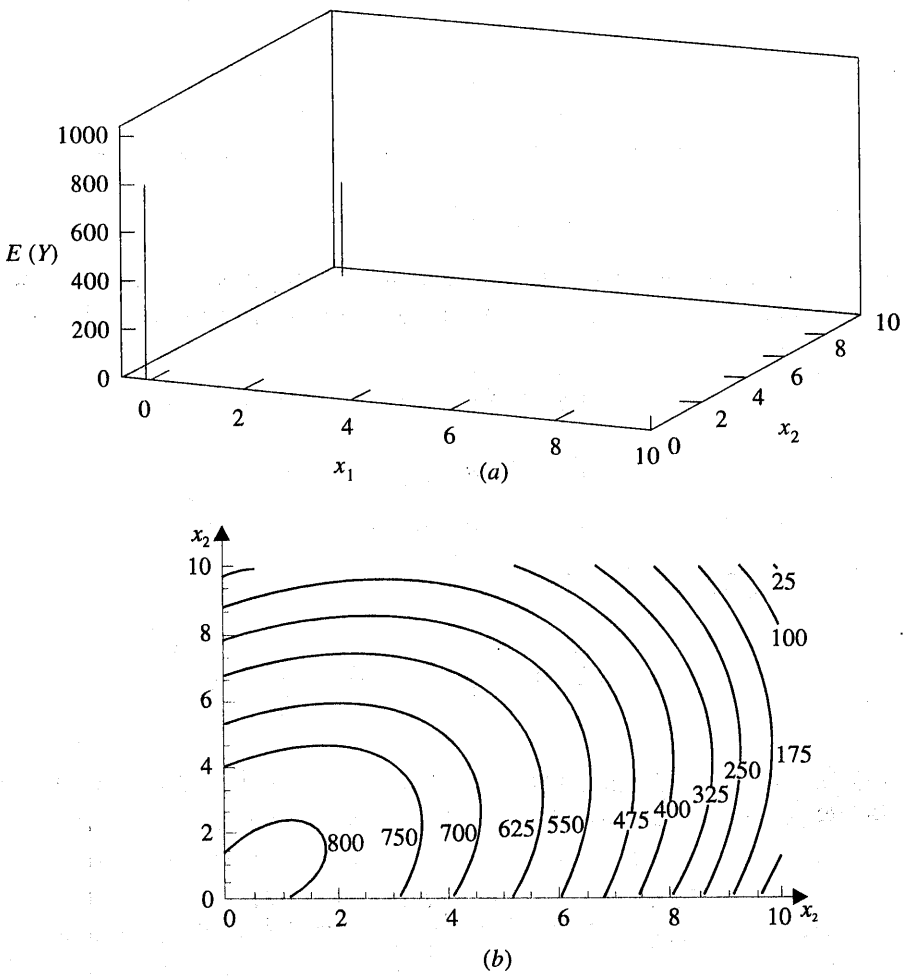


Figura 11-3 ***a***) Gráfica tridimensional del modelo de regresión $E(Y) = 800 + 10x_1 + 7x_2 - 8.5x_1^2 - 5x_2^2 + 4x_1x_2$. ***b***) La gráfica de contorno.

Si se hace $x_3 = x_1^2, x_4 = x_2^2, x_5 = x_1x_2, \beta_3 = \beta_{11}, \beta_4 = \beta_{22}$, y $\beta_5 = \beta_{12}$, entonces la ecuación 11-6 puede escribirse como un modelo de regresión lineal múltiple como sigue:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \varepsilon$$

En la figura 11-3*a* y *b* se muestra la gráfica tridimensional y la gráfica de contorno (curvas de nivel) correspondiente para

$$E(Y) = 800 + 10x_1 + 7x_2 - 8.5x_1^2 - 5x_2^2 + 4x_1x_2$$

Estas gráficas indican que el cambio esperado en Y cuando x_1 cambia una unidad (por ejemplo) es una función *tanto* de x_1 *como* de x_2. El término cuadrático y el de interacción de este modelo producen una función en forma de montículo. Dependiendo de los valores de los coeficientes de regresión, el modelo de segundo orden con interacción puede adoptar una amplia variedad de formas; por tanto, se trata de un modelo de regresión muy flexible.

11-2 ESTIMACIÓN DE LOS PARÁMETROS POR MÍNIMOS CUADRADOS

Puede usarse el método de mínimos cuadrados para estimar los coeficientes de regresión en el modelo de regresión múltiple, ecuación 11-2. Suponga que se cuenta con $n > k$ observaciones, y sea que x_{ij} denote la observación o nivel i de la variable x_j. Las observaciones son

$$(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i), \qquad i = 1, 2, \ldots, n \quad \text{y} \quad n > k$$

Se acostumbra presentar los datos de una regresión múltiple en una tabla como la tabla 11-1.

Cada observación $(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i)$ satisface el modelo de la ecuación 11-2, o

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \\ &= \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i \qquad i = 1, 2, \ldots, n \end{aligned} \tag{11-7}$$

La función de mínimos cuadrados es

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2 \tag{11-8}$$

La función L debe minimizarse con respecto a $\beta_0, \beta_1, \ldots, \beta_k$. Las **estimaciones de mínimos cuadrados** de $\beta_0, \beta_1, \ldots, \beta_k$ deben satisfacer

$$\left. \frac{\partial L}{\partial \beta_0} \right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) = 0 \tag{11-9a}$$

y

$$\left. \frac{\partial L}{\partial \beta_j} \right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) x_{ij} = 0 \qquad j = 1, 2 \ldots, k \tag{11-9b}$$

Tabla 11-1 Datos de una regresión lineal múltiple

y	x_1	x_2	$\cdots$	x_k
y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}

Al simplificar la ecuación 11-9 se obtienen las **ecuaciones normales de mínimos cuadrados**

$$
\begin{aligned}
n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \ldots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} &= \sum_{i=1}^{n} y_i \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1} x_{i2} + \ldots + \hat{\beta}_k \sum_{i=1}^{n} x_{i1} x_{ik} &= \sum_{i=1}^{n} x_{i1} y_i \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_1 \sum_{i=1}^{n} x_{ik} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ik} x_{i2} + \ldots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 &= \sum_{i=1}^{n} x_{ik} y_i \qquad (11\text{-}10)
\end{aligned}
$$

Obsérvese que hay $p = k + 1$ ecuaciones normales, una para cada uno de los coeficientes de regresión desconocidos. Las soluciones de las ecuaciones normales serán los estimadores de mínimos cuadrados de los coeficientes de regresión, $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$. Las soluciones de las ecuaciones normales pueden encontrarse con cualquier método apropiado para resolver un sistema de ecuaciones lineales.

EJEMPLO 11-1

En el capítulo 1 se usaron los datos de la resistencia al desprendimiento de un alambre adherido en un proceso de manufactura de semiconductores, de la longitud del alambre y de la altura de la matriz para ilustrar la construcción de un modelo empírico. Se usarán los mismos datos, los cuales se repiten por comodidad en la tabla 11-2, y se mostrarán los detalles de la estimación de los parámetros del modelo. En la figura 1-8 se presenta un diagrama de dispersión tridimensional de los datos.

Específicamente, se ajustará el modelo de regresión lineal múltiple

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

donde Y = resistencia al desprendimiento, x_1 = longitud del alambre y x_2 = altura de la matriz. A partir de los datos de la tabla 11-2 se calcula

$$n = 25, \quad \sum_{i=1}^{25} y_i = 725.82$$

$$\sum_{i=1}^{25} x_{i1} = 206, \quad \sum_{i=1}^{25} x_{i2} = 8\,294$$

$$\sum_{i=1}^{25} x_{i1}^2 = 2\,396, \quad \sum_{i=1}^{25} x_{i2}^2 = 3\,531\,848$$

$$\sum_{i=1}^{25} x_{i1} x_{i2} = 77\,177, \quad \sum_{i=1}^{25} x_{i1} y_i = 8\,008.37, \quad \sum_{i=1}^{25} x_{i2} y_i = 274\,811.31$$

Tabla 11-2 Datos del alambre adherido del alambre del ejemplo 11-1

Número de observación	Resistencia al desprendimiento y	Longitud del alambre x_1	Altura de la matriz x_2
1	9.95	2	50
2	24.45	8	110
3	31.75	11	120
4	35.00	10	550
5	25.02	8	295
6	16.86	4	200
7	14.38	2	375
8	9.60	2	52
9	24.35	9	100
10	27.50	8	300
11	17.08	4	412
12	37.00	11	400
13	41.95	12	500
14	11.66	2	360
15	21.65	4	205
16	17.89	4	400
17	69.00	20	600
18	10.30	1	585
19	34.93	10	540
20	46.59	15	250
21	44.88	15	290
22	54.12	16	510
23	56.23	17	590
24	22.13	6	100
25	21.15	5	400

Para el modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, las ecuaciones normales 11-10 son

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1} x_{i2} = \sum_{i=1}^{n} x_{i1} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_{i2} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} x_{i2} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2}^2 = \sum_{i=1}^{n} x_{i2} y_i$$

Al sustituir las sumatorias calculadas en las ecuaciones normales, se obtiene

$$\begin{aligned} 25\hat{\beta}_0 + 206\hat{\beta}_1 + 8\,294\hat{\beta}_2 &= 725.82 \\ 206\hat{\beta}_0 + 2\,396\hat{\beta}_1 + 77\,177\hat{\beta}_2 &= 8\,008.37 \\ 8\,294\hat{\beta}_0 + 77\,177\hat{\beta}_1 + 3\,531\,848\hat{\beta}_2 &= 274\,811.31 \end{aligned}$$

La solución de este sistema de ecuaciones es

$$\hat{\beta}_0 = 2.26379, \quad \hat{\beta}_1 = 2.74427, \quad \hat{\beta}_2 = 0.01253$$

Por lo tanto, la ecuación de regresión ajustada es

$$\hat{y} = 2.26379 + 2.74427x_1 + 0.01253x_2$$

Esta ecuación puede usarse para predecir la resistencia al desprendimiento para pares de valores de las variables regresoras de la longitud del alambre (x_1) y la altura de la matriz (x_2). Se trata en esencia del mismo modelo de regresión dado en la ecuación 1-12, sección 1-3. En la figura 1-9 se presenta una gráfica de los valores predichos $\hat{y}$ generados por esta ecuación.

11-3 ENFOQUE MATRICIAL DE LA REGRESIÓN LINEAL MÚLTIPLE

Al ajustar un modelo de regresión múltiple, es mucho más conveniente expresar las operaciones matemáticas utilizando la notación matricial. Suponga que hay k variables regresoras y n observaciones $(x_{i1}, x_{i2}, ..., x_{ik}, y_i)$, $i = 1, 2, ..., n$, y que el modelo que relaciona las variables regresoras con la variable de respuesta es

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \quad i = 1, 2, ..., n$$

Este modelo es un sistema de n ecuaciones que en notación matricial puede expresarse como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{11-11}$$

donde

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \text{y} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

En general, $\mathbf{y}$ es un vector $(n \times 1)$ de las observaciones, $\mathbf{X}$ es una matriz $(n \times p)$ de los niveles de las variables independientes, $\boldsymbol{\beta}$ es un vector $(p \times 1)$ de los coeficientes de regresión, y $\boldsymbol{\varepsilon}$ es un vector $(n \times 1)$ de los errores aleatorios.

Quiere encontrarse el vector de los estimadores de mínimos cuadrados, $\hat{\boldsymbol{\beta}}$, que minimice

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

El estimador de mínimos cuadrados $\hat{\boldsymbol{\beta}}$ es la solución para $\boldsymbol{\beta}$ en las ecuaciones

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \mathbf{0}$$

No se presentarán los detalles para sacar las derivadas anteriores; sin embargo, las ecuaciones resultantes que deben resolverse son

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \tag{11-12}$$

Las ecuaciones 11-12 son las ecuaciones normales de mínimos cuadrados en forma matricial. Son idénticas a la forma escalar de las ecuaciones normales que se presentaron antes en las ecuaciones 11-10. Para resolver las ecuaciones normales, se multiplican ambos miembros de las ecuaciones 11-12 por la inversa de $\mathbf{X}'\mathbf{X}$. Por lo tanto, la estimación de mínimos cuadrados de $\boldsymbol{\beta}$ es

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{11-13}$$

Obsérvese que hay $p = k + 1$ ecuaciones normales en $p = k + 1$ incógnitas (los valores de $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$). Además, la matriz $\mathbf{X}'\mathbf{X}$ con frecuencia es no singular, como se supuso arriba, por lo que los métodos descritos en los libros de texto sobre determinantes y matrices para invertir una matriz pueden usarse para encontrar $(\mathbf{X}'\mathbf{X})^{-1}$. En la práctica, los cálculos de una regresión múltiple se hacen con una computadora.

Es fácil ver que la forma matricial de las ecuaciones normales es idéntica a la forma escalar. Al escribir la ecuación 11-12 en detalle, se obtiene

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} & \cdots & \sum_{i=1}^{n} x_{i1}x_{ik} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{ik}x_{i1} & \sum_{i=1}^{n} x_{ik}x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik}y_i \end{bmatrix}$$

Si se realiza la multiplicación de matrices indicada, se llegará a la forma escalar de las ecuaciones normales (es decir, a la ecuación 11-10). En esta forma es sencillo ver que $\mathbf{X}'\mathbf{X}$ es una matriz simétrica $(p \times p)$ y que $\mathbf{X}'\mathbf{y}$ es un vector columna $(p \times 1)$. Nótese la estructura especial de la matriz $\mathbf{X}'\mathbf{X}$. Los elementos de la diagonal de $\mathbf{X}'\mathbf{X}$ son las sumas de los cuadrados de los elementos de las columnas de $\mathbf{X}$, y los elementos que están fuera de la diagonal son las sumas de los productos cruzados de los elementos de las columnas de $\mathbf{X}$. Además, obsérvese que los elementos de $\mathbf{X}'\mathbf{y}$ son las sumas de los productos cruzados de las columnas de $\mathbf{X}$ y las observaciones $\{y_i\}$.

El modelo de regresión ajustado es

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \qquad i = 1, 2, \ldots, n \tag{11-14}$$

En notación matricial, el modelo ajustado es

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$$

La diferencia entre la observación y_i y el valor ajustado $\hat{y}_i$ es un **residual**, es decir, $e_i = y_i - \hat{y}_i$. El vector $(n \times 1)$ de los residuales se denota por

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \tag{11-15}$$

EJEMPLO 11-2

En el ejemplo 11-1 se ilustró el ajuste del modelo de regresión múltiple

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

donde y es la resistencia al desprendimiento observada del alambre adherido, x_1 es la longitud del alambre, y x_2 es la altura de la matriz. En la tabla 11-2 se muestran las 25 observaciones. Se usará ahora el enfoque matricial para ajustar el modelo de regresión anterior a estos datos. La matriz $\mathbf{X}$ y el vector $\mathbf{y}$ para este modelo son

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 50 \\ 1 & 8 & 110 \\ 1 & 11 & 120 \\ 1 & 10 & 550 \\ 1 & 8 & 295 \\ 1 & 4 & 200 \\ 1 & 2 & 375 \\ 1 & 2 & 52 \\ 1 & 9 & 100 \\ 1 & 8 & 300 \\ 1 & 4 & 412 \\ 1 & 11 & 400 \\ 1 & 12 & 500 \\ 1 & 2 & 360 \\ 1 & 4 & 205 \\ 1 & 4 & 400 \\ 1 & 20 & 600 \\ 1 & 1 & 585 \\ 1 & 10 & 540 \\ 1 & 15 & 250 \\ 1 & 15 & 290 \\ 1 & 16 & 510 \\ 1 & 17 & 590 \\ 1 & 6 & 100 \\ 1 & 5 & 400 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 9.95 \\ 24.45 \\ 31.75 \\ 35.00 \\ 25.02 \\ 16.86 \\ 14.38 \\ 9.60 \\ 24.35 \\ 27.50 \\ 17.08 \\ 37.00 \\ 41.95 \\ 11.66 \\ 21.65 \\ 17.89 \\ 69.00 \\ 10.30 \\ 34.93 \\ 46.59 \\ 44.88 \\ 54.12 \\ 56.63 \\ 22.13 \\ 21.15 \end{bmatrix}$$

La matriz $\mathbf{X'X}$ es

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 2 & 8 & \cdots & 5 \\ 50 & 110 & \cdots & 400 \end{bmatrix} \begin{bmatrix} 1 & 2 & 50 \\ 1 & 8 & 110 \\ \vdots & \vdots & \vdots \\ 1 & 5 & 400 \end{bmatrix}$$

$$\begin{bmatrix} 25 & 206 & 8\,294 \\ 206 & 2\,396 & 77\,177 \\ 8\,294 & 77\,177 & 3\,531\,848 \end{bmatrix}$$

y el vector $\mathbf{X'y}$ es

$$\mathbf{X'y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 2 & 8 & \cdots & 5 \\ 50 & 110 & \cdots & 400 \end{bmatrix} \begin{bmatrix} 9.95 \\ 24.45 \\ \vdots \\ 21.15 \end{bmatrix} = \begin{bmatrix} 725.82 \\ 8\,008.37 \\ 274\,811.31 \end{bmatrix}$$

Las estimaciones de mínimos cuadrados se encuentran con la ecuación 11-13 como

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{y}$$

o bien

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 25 & 206 & 8\,294 \\ 206 & 2\,396 & 77\,177 \\ 8\,294 & 77\,177 & 3\,531\,848 \end{bmatrix}^{-1} \begin{bmatrix} 725.82 \\ 8\,008.37 \\ 274\,811.31 \end{bmatrix}$$

$$= \begin{bmatrix} 0.21653 & -0.007491 & -0.000340 \\ -0.007491 & 0.001671 & -0.000019 \\ -0.000340 & -0.000019 & -0.0000015 \end{bmatrix} \begin{bmatrix} 725.82 \\ 8\,008.47 \\ 274\,811.31 \end{bmatrix}$$

$$= \begin{bmatrix} 2.26379143 \\ 2.74426964 \\ 0.01252781 \end{bmatrix}$$

Por lo tanto, el modelo de regresión ajustado con los coeficientes de regresión redondeados a cinco cifras decimales es

$$\hat{y} = 2.26379 + 2.74427x_1 + 0.01253x_2$$

Estos resultados son idénticos a los obtenidos en el ejemplo 11-1.

Este modelo de regresión puede usarse para predecir los valores de la resistencia al desprendimiento para varios valores de la longitud del alambre (x_1) y la altura de la matriz (x_2). También es posible obtener los **valores ajustados** $\hat{y}_i$, sustituyendo cada observación (x_{i1}, x_{i2}), $i = 1, 2, ..., n$, en la ecuación. Por ejemplo, la primera observación tiene $x_{11} = 2$ y $x_{12} = 50$, y el valor ajustado es

$$\begin{aligned} \hat{y}_1 &= 2.26379 + 2.74427x_{11} + 0.01253x_{12} \\ &= 2.26379 + 2.74427(2) + 0.01253(50) \\ &= 8.38 \end{aligned}$$

El valor observado correspondiente es $y_1 = 9.95$. El *residual* correspondiente de la primera observación es

$$\begin{aligned} e_1 &= y_1 - \hat{y}_1 \\ &= 9.95 - 8.38 \\ &= 1.57 \end{aligned}$$

En la tabla 11-3 se muestran los 25 valores ajustados $\hat{y}_i$ y los residuales correspondientes. Los valores ajustados y los residuales se calcularon con la misma precisión que los datos originales.

Tabla 11-3 Observaciones, valores ajustados y residuales del ejemplo 11-2

Número de observación	y_i	$\hat{y}_i$	$e_i = y_i - \hat{y}_i$
1	9.95	8.38	1.57
2	24.45	25.60	–1.15
3	31.75	33.95	–2.20
4	35.00	36.60	– 1.60
5	25.02	27.91	– 2.89
6	16.86	15.75	1.11
7	14.38	12.45	1.93
8	9.60	8.40	1.20
9	24.35	28.21	– 3.86
10	27.50	27.98	– 0.48
11	17.08	18.40	– 1.32
12	37.00	37.46	– 0.46
13	41.95	41.46	0.49
14	11.66	12.26	– 0.60
15	21.65	15.81	5.84
16	17.89	18.25	– 0.36
17	69.00	64.67	4.33
18	10.30	12.34	– 2.04
19	34.93	36.47	– 1.54
20	46.59	46.56	– 0.03
21	44.88	47.06	– 2.18
22	54.12	52.56	1.56
23	56.63	56.31	0.32
24	22.13	19.98	2.15
25	21.15	21.00	0.15

Casi siempre se usan computadoras para ajustar los modelos de regresión múltiple. En la tabla 11-4 se presenta una parte de la salida del programa Minitab para los datos de la resistencia al desprendimiento del alambre adherido de los ejemplos 11-1 y 11-2. Obsérvese que la parte superior de la tabla contiene las estimaciones numéricas de los coeficientes de regresión. La computadora también calcula otras cantidades que reflejan información importante acerca del modelo de regresión. En secciones posteriores se definirán y explicarán las cantidades de esta salida de computadora.

Tabla 11-4 Salida del programa Minitab para los datos de la resistencia al desprendimiento del alambre adherido

Análisis de regresión

La ecuación de regresión es
$y = 2.26 + 2.74x_1 + 0.0125x_2$

	Coeficiente	Desviación estándar	T	P
Constante de predicción	2.264	1.060	2.14	0.044
x_1	2.74427	0.09352	29.34	0.000
x_2	0.012528	0.002798	4.48	0.000

S = 2.288 R. cuadrada = 98.1% R. cuadrada (ajustada) = 97.9%

Análisis de varianza

Fuente	GL	SS	MS	F	P
Regresión	2	5 990.8	2 995.4	572.17	0.000
Error	22	115.2	5.2		
Total	24	6 105.9			

Fuente	GL	Seq SS
x_1	1	5 885.9
x_2	1	104.9

EJERCICIOS DE LAS SECCIONES 11-2 Y 11-3

11-1. Se realizó un estudio para investigar la relación de la resistencia al corte del terreno (y) con la profundidad en pies (x_1) y el contenido de humedad (x_2). Se hicieron 10 observaciones, obteniéndose las siguientes cantidades resumidas: $n = 10$, $\Sigma x_{i1} = 223$, $\Sigma x_{i2} = 553$, $\Sigma y_i = 1\,916$, $\Sigma x_{i1}^2 = 5\,200.9$, $\Sigma x_{i2}^2 = 31\,729$, $\Sigma x_{i1}x_{i2} = 12\,352$, $\Sigma x_{i1}y_i = 43\,550.8$, $\Sigma x_{i2}y_i = 104\,736.8$ y $\Sigma y_i^2 = 371\,595.6$.

a) Establezca las ecuaciones normales de mínimos cuadrados para el modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$.

b) Estime los parámetros del modelo del inciso a).

c) ¿Cuál es la resistencia predicha cuando $x_1 = 18$ pies y $x_2 = 43\%$?

11-2. Debe desarrollarse un modelo de regresión para predecir la capacidad del suelo para absorber contaminantes químicos. Se hacen 10 observaciones del índice de absorción (y) de un suelo y de dos regresores: x_1 = cantidad de mineral de hierro extraíble, y x_2 = cantidad de bauxita. Quiere ajustarse el modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$. Algunas de las cantidades necesarias son:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 1.17991 & -7.30982\text{ E-3} & 7.3006\text{ E-4} \\ -7.30982\text{ E-3} & 7.9799\text{ E-5} & -1.23713\text{ E-4} \\ 7.3006\text{ E-4} & -1.23713\text{ E-4} & -4.6576\text{ E-4} \end{bmatrix}, \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 220 \\ 36\,768 \\ 9\,965 \end{bmatrix}$$

a) Estime los coeficientes de regresión del modelo especificado arriba.

b) ¿Cuál es el valor predicho del índice de absorción y cuando $x_1 = 200$ y $x_2 = 50$?

11-3. Un ingeniero químico está investigando la manera en que la cantidad de conversión de un producto de una materia prima (y) depende de la temperatura de reacción (x_1) y del tiempo de reacción (x_2). El ingeniero ha desarrollado los siguientes modelos de regresión:

1. $\hat{y} = 100 + 2x_1 + 4x_2$

2. $\hat{y} = 95 + 1.5x_1 + 3x_2 + 2x_1x_2$

Ambos modelos se han construido en el rango $0.5 \le x_2 \le 10$.

a) ¿Cuál es el valor predicho de la conversión cuando $x_2 = 2$? Repita este cálculo para $x_2 = 8$. Trace una gráfica de los valores predichos para ambos modelos de la conversión. Comente el efecto del término de interacción del modelo 2.

b) Encuentre el cambio esperado en la conversión promedio para un cambio unitario en la temperatura x_1 del modelo 1 cuando $x_2 = 5$. ¿Esta cantidad depende del valor específico del tiempo de reacción seleccionado? ¿Por qué?

c) Encuentre el cambio esperado en la conversión promedio para un cambio unitario en la temperatura x_1 del modelo 2 cuando $x_2 = 5$. Repita este cálculo para $x_2 = 2$ y $x_2 = 8$. ¿El resultado depende del valor seleccionado para x_2? ¿Por qué?

11-4. Los datos de la tabla de la página 499 son las estadísticas del desempeño de los equipos de la Liga Nacional de Futbol Americano (NFL, por sus siglas en inglés) en 1976 (*Fuente: The Sporting News*).

a) Ajuste un modelo de regresión múltiple que relacione el número de juegos ganados con las yardas por pase de los equipos (x_2), el porcentaje de jugadas por carrera (x_7) y las yardas por carrera de los oponentes (x_8).

b) Use el modelo para predecir el número de juegos ganados cuando $x_2 = 2\,000$ yardas, $x_7 = 60\%$ y $x_8 = 1\,800$ yardas.

11-5. En la tabla de la página 500 se presenta el rendimiento de gasolina (millas/galón) de 25 automóviles (*Fuente: Motor Trend*).

a) Ajuste un modelo de regresión múltiple que relacione el rendimiento de gasolina con el cilindraje del motor (x_1) y el número de gargantas del carburador (x_6).

b) Use el modelo desarrollado en el inciso a) para predecir el rendimiento de gasolina cuando el motor tiene un desplazamiento $x_1 = 300$ pulg^3 y un carburador de cuatro gargantas.

11-6. Se piensa que la energía eléctrica consumida mensualmente por una planta química se relaciona con la temperatura ambiente promedio (x_1), el número de días laborales del mes (x_2), la pureza promedio del producto (x_3) y las toneladas del producto producidas (x_4). Se cuenta con los datos del último año, los cuales se presentan en la tabla siguiente:

y	x_1	x_2	x_3	x_4
240	25	24	91	100
236	31	21	90	95
290	45	24	88	110
274	60	25	87	88
301	65	25	91	94
316	72	26	94	99
300	80	25	87	97
296	84	25	86	96
267	75	24	88	110
276	60	25	91	105
288	50	25	90	100
261	38	23	89	98

a) Ajuste un modelo de regresión lineal múltiple a estos datos.

b) Prediga el consumo de electricidad para un mes en el que $x_1 = 75°F$, $x_2 = 24$ días, $x_3 = 90\%$ y $x_4 = 98$ toneladas.

Desempeño de los equipos de la Liga Nacional de Futbol Americano (NFL) en 1976

Equipo	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
Washington	10	2 113	1 985	38.9	64.7	+4	868	59.7	2 205	1 917
Minnesota	11	2 003	2 855	38.8	61.3	+3	615	55.0	2 096	1 575
Nueva Inglaterra	11	2 957	1 737	40.1	60.0	+14	914	65.6	1 847	2 175
Oakland	13	2 285	2 905	41.6	45.3	–4	957	61.4	1 903	2 476
Pittsburgh	10	2 971	1 666	39.2	53.8	+15	836	66.1	1 457	1 866
Baltimore	11	2 309	2 927	39.7	74.1	+8	786	61.0	1 848	2 339
Los Ángeles	10	2 528	2 341	38.1	65.4	+12	754	66.1	1 564	2 092
Dallas	11	2 147	2 737	37.0	78.3	–1	797	58.9	2 476	2 254
Atlanta	4	1 689	1 414	42.1	47.6	–3	714	57.0	2 577	2 001
Búffalo	2	2 566	1 838	42.3	54.2	–1	797	58.9	2 476	2 254
Chicago	7	2 363	1 480	37.3	48.0	+19	984	68.5	1 984	2 217
Cincinnati	10	2 109	2 191	39.5	51.9	+6	819	59.2	1 901	1 686
Cleveland	9	2 295	2 229	37.4	53.6	–5	1 037	58.8	1 761	2 032
Denver	9	1 932	2 204	35.1	71.4	+3	986	58.6	1 709	2 025
Detroit	6	2 213	2 140	38.8	58.3	+6	819	59.2	1 901	1 686
Green Bay	5	1 722	1 730	36.6	52.6	–19	791	54.4	2 288	1 835
Houston	5	1 498	2 072	35.3	59.3	–5	776	49.6	2 072	1 914
Kansas City	5	1 873	2 929	41.1	55.3	+10	789	54.3	2 861	2 496
Miami	6	2 118	2 268	38.6	69.6	+6	582	58.7	2 411	2 670
Nueva Orleans	4	1 775	1 983	39.3	78.3	+7	901	51.7	2 289	2 202
Gigantes de Nueva York	3	1 904	1 792	39.7	38.1	–9	734	61.9	2 203	1 988
Jets de Nueva York	3	1 929	1 606	39.7	68.8	–21	627	52.7	2 592	2 324
Filadelfia	4	2 080	1 492	35.5	68.8	–8	722	57.8	2 053	2 550
San Luis	10	2 301	2 835	35.3	74.1	+2	683	59.7	1 979	2 110
San Diego	6	2 040	2 416	38.7	50.0	0	576	54.9	2 048	2 628
San Francisco	8	2 447	1 638	39.9	57.1	–8	848	65.3	1 786	1 776
Seattle	2	1 416	2 649	37.4	56.3	–22	684	43.8	2 876	2 524
Tampa Bay	0	1 503	1 503	39.3	47.0	–9	875	53.5	2 560	2 241

y: juegos ganados (de 14 juegos de la temporada)
x_1: yardas por carrera (temporada)
x_2: yardas por pase (temporada)
x_3: yardas del pateador (yardas/patada)
x_4: porcentaje de goles de campo (goles de campo conseguidos/goles de campo intentados; temporada)
x_5: diferencial de balones sueltos (balones sueltos recuperados – balones sueltos perdidos)
x_6: yardas perdidas por castigo (temporada)
x_7: porcentaje de jugadas por carrera (jugadas por carrera/jugadas totales)
x_8: yardas por carrera de los oponentes (temporada)
x_9: yardas por pase de los oponentes (temporada)

Rendimiento de gasolina de 25 automóviles

Automóvil	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
Apollo	18.90	350	165	260	8.0:1	2.56:1	4	3	200.3	69.9	3 910	A
Nova	20.00	250	105	185	8.25:1	2.73:1	1	3	196.7	72.2	3 510	A
Monarch	18.25	351	143	255	8.0:1	3.00:1	2	3	199.9	74.0	3 890	A
Duster	20.07	225	95	170	8.4:1	2.76:1	1	3	194.1	71.8	3 365	M
Jenson Conv.	11.2	440	215	330	8.2:1	2.88:1	4	3	184.5	69	4 215	A
Skyhawk	22.12	231	110	175	8.0:1	2.56:1	2	3	179.3	65.4	3 020	A
Sciroeco	34.70	89.7	70	81	8.2:1	3.90:1	2	4	155.7	64	1 905	M
Corolla SR-5	30.40	96.9	75	83	9.0:1	4.30:1	2	5	165.2	65	2 320	M
Camaro	16.50	350	155	250	8.5:1	3.08:1	4	3	195.4	74.4	3 885	A
Datsun B210	36.50	85.3	80	83	8.5:1	3.89:1	2	4	160.6	62.2	2 009	M
Capri 11	21.50	171	109	146	8.2:1	3.22:1	2	4	170.4	66.9	2 655	M
Pacer	19.70	258	110	195	8.0:1	3.08:1	1	3	171.5	77	3 375	A
Granada	17.80	302	129	220	8.0:1	3.0:1	2	3	199.9	74	3 890	A
Eldorado	14.39	500	190	360	8.5:1	2.73:1	4	3	224.1	79.8	5 290	A
Imperial	14.89	440	215	330	8.2:1	2.71:1	4	3	231.0	79.7	5 185	A
Nova LN	17.80	350	155	250	8.5:1	3.08:1	4	3	196.7	72.2	3 910	A
Starfire	23.54	231	110	175	8.0:1	2.56:1	2	3	179.3	65.4	3 050	A
Cordoba	21.47	360	180	290	8.4:1	2.45:1	2	3	214.2	76.3	4 250	A
Trans Am	16.59	400	185	NA	7.6:1	3.08:1	4	3	196	73	3 850	A
Corolla E-5	31.90	96.9	75	83	9.0:1	4.30:1	2	5	165.2	61.8	2 275	M
Mark IV	13.27	460	223	366	8.0:1	3.00:1	4	3	228	79.8	5 430	A
Celica GT	23.90	133.6	96	120	8.4:1	3.91:1	2	5	171.5	63.4	2 535	M
Charger SE	19.73	318	140	255	8.5:1	2.71:1	2	3	215.3	76.3	4 370	A
Cougar	13.90	351	148	243	8.0:1	3.25:1	2	3	215.5	78.5	4 540	A
Corvette	16.50	350	165	255	8.5:1	2.73:1	4	3	185.2	69	3 660	A

y: millas/galón
x_1: desplazamiento (pulgadas cúbicas)
x_2: caballos de fuerza (libras-pie)
x_3: torque (libras-pie)
x_4: relación de compresión
x_5: relación del eje trasero
x_6: carburador (gargantas)
x_7: número de velocidades
x_8: longitud total (pulgadas)
x_9: ancho (pulgadas)
x_{10}: peso (libras)
x_{11}: tipo de transmisión (A-automática; M-manual)

11-7. Se realizó un estudio del desgaste de un rodamiento y y su relación con x_1 = viscosidad del aceite y x_2 = carga. Se obtuvieron los siguientes datos:

y	x_1	x_2
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

a) Ajuste un modelo de regresión lineal múltiple a estos datos.

b) Use el modelo para predecir el desgaste cuando $x_1 = 25$ y $x_2 = 1\ 000$.

c) Ajuste un modelo de regresión lineal múltiple con un término de interacción a estos datos.

d) Use el modelo del inciso c) para predecir el desgaste cuando $x_1 = 25$ y $x_2 = 1\ 000$. Compare esta predicción con el valor predicho en el inciso b).

11-8. La resistencia al desprendimiento de un alambre adherido es una característica importante. En la tabla siguiente se ofrece información sobre la resistencia al desprendimiento (y), la altura de la matriz (x_1), la altura del poste (x_2), altura del amarre (x_3), longitud del alambre (x_4), anchura del amarre en la matriz (x_5) y anchura del amarre en el poste (x_6).

y	x_1	x_2	x_3	x_4	x_5	x_6
8.0	5.2	19.6	29.6	94.9	2.1	2.3
8.3	5.2	19.8	32.4	89.7	2.1	1.8
8.5	5.8	19.6	31.0	96.2	2.0	2.0
8.8	6.4	19.4	32.4	95.6	2.2	2.1
9.0	5.8	18.6	28.6	86.5	2.0	1.8
9.3	5.2	18.8	30.6	84.5	2.1	2.1
9.3	5.6	20.4	32.4	88.8	2.2	1.9
9.5	6.0	19.0	32.6	85.7	2.1	1.9
9.8	5.2	20.8	32.2	93.6	2.3	2.1
10.0	5.8	19.9	31.8	86.0	2.1	1.8
10.3	6.4	18.0	32.6	87.1	2.0	1.6
10.5	6.0	20.6	33.4	93.1	2.1	2.1
10.8	6.2	20.2	31.8	83.4	2.2	2.1
11.0	6.2	20.2	32.4	94.5	2.1	1.9
11.3	6.2	19.2	31.4	83.4	1.9	1.8
11.5	5.6	17.0	33.2	85.2	2.1	2.1
11.8	6.0	19.8	35.4	84.1	2.0	1.8
12.3	5.8	18.8	34.0	86.9	2.1	1.8
12.5	5.6	18.6	34.2	83.0	1.9	2.0

a) Ajuste un modelo de regresión lineal múltiple usando x_2, x_3, x_4 y x_5 como variables regresoras.

b) Use el modelo del inciso a) para predecir la resistencia al desprendimiento cuando $x_2 = 20$, $x_3 = 30$, $x_4 = 90$ y $x_5 = 2.0$.

11-9. Un ingeniero de una compañía de semiconductores quiere modelar la relación entre la HFE del dispositivo (y) y tres parámetros: RS del emisor (x_1), RS de la base (x_2) y RS emisor-base (x_3). Los datos se muestran en la tabla siguiente.

x_1 RS del emisor	x_2 RS de la base	x_3 RS e-b	y HFE-1M-5V
14.620	226.00	7.000	128.40
15.630	220.00	3.375	52.62
14.620	217.40	6.375	113.90
15.000	220.00	6.000	98.01
14.500	226.50	7.625	139.90
15.250	224.10	6.000	102.60
16.120	220.50	3.375	48.14
15.130	223.50	6.125	109.60
15.500	217.60	5.000	82.68
15.130	228.50	6.625	112.60
15.500	230.20	5.750	97.52
16.120	226.50	3.750	59.06
15.130	226.60	6.125	111.80
15.630	225.60	5.375	89.09
15.380	229.70	5.875	101.00
14.380	234.00	8.875	171.90
15.500	230.00	4.000	66.80
14.250	224.30	8.000	157.10
14.500	240.50	10.870	208.40
14.620	223.70	7.375	133.40

a) Ajuste un modelo de regresión lineal múltiple a estos datos.

b) Prediga la HFE cuando $x_1 = 14.5$, $x_2 = 220$ y $x_3 = 5.0$.

11-10. Es común el uso de tratamientos térmicos para carburar piezas metálicas, tales como engranes. El espesor de la capa carburada se considera una característica determinante del engrane y contribuye a la confiabilidad global de la pieza. Debido a la naturaleza crítica de esta característica, se realizan dos pruebas de laboratorio diferentes para cada carga del horno. Una prueba se hace en un perno de prueba que acompaña a cada carga. La otra prueba es de carácter destructivo, en la que se hace un corte transversal en una pieza original. Esta prueba incluye un análisis de carbono en la superficie tanto del paso del engrane (la cara superior del diente del engrane) como de la base del engrane (el espacio entre los dientes del engrane). Los datos siguientes son los resultados del análisis de carbono en el paso de 32 engranes.

Las variables regresoras son la temperatura del horno (TEMP), la concentración de carbono y la duración del ciclo de carburación (SOAKPCT, SOAKTIME), y la concentración de carbono y la duración del ciclo difusión (DIFFPCT, DIFFTIME).

a) Ajuste un modelo de regresión lineal que relacione los resultados del análisis de carbono en el paso (PITCH) con las cinco variables regresoras.

b) Use el modelo del inciso a) para predecir el paso cuando TEMP = 1 650, SOAKTIME = 1.00, SOAKPCT = 1.10, DIFFTIME = 1.00 y DIFFPCT = 0.80.

11-11. Se obtuvieron las estadísticas de los 21 equipos de la Liga Nacional de Hockey sobre

TEMP	SOAKTIME	SOAKPCT	DIFFTIME	DIFFPCT	PITCH
1650	0.58	1.10	0.25	0.90	0.013
1650	0.66	1.10	0.33	0.90	0.016
1650	0.66	1.10	0.33	0.90	0.015
1650	0.66	1.10	0.33	0.95	0.016
1600	0.66	1.15	0.33	1.00	0.015
1600	0.66	1.15	0.33	1.00	0.016
1650	1.00	1.10	0.50	0.80	0.014
1650	1.17	1.10	0.58	0.80	0.021
1650	1.17	1.10	0.58	0.80	0.018
1650	1.17	1.10	0.58	0.80	0.019
1650	1.17	1.10	0.58	0.90	0.021
1650	1.17	1.10	0.58	0.90	0.019
1650	1.17	1.15	0.58	0.90	0.021
1650	1.20	1.15	1.10	0.80	0.025
1650	2.00	1.15	1.00	0.80	0.025
1650	2.00	1.10	1.10	0.80	0.026
1650	2.20	1.10	1.10	0.80	0.024
1650	2.20	1.10	1.10	0.80	0.025
1650	2.20	1.15	1.10	0.80	0.024
1650	2.20	1.10	1.10	0.90	0.025
1650	2.20	1.10	1.10	0.90	0.027
1650	2.20	1.10	1.50	0.90	0.026
1650	3.00	1.15	1.50	0.80	0.029
1650	3.00	1.10	1.50	0.70	0.030
1650	3.00	1.10	1.50	0.75	0.028
1650	3.00	1.15	1.66	0.85	0.032
1650	3.33	1.10	1.50	0.80	0.033
1700	4.00	1.10	1.50	0.70	0.039
1650	4.00	1.10	1.50	0.70	0.040
1650	4.00	1.15	1.50	0.85	0.035
1700	12.50	1.00	1.50	0.70	0.056
1700	18.50	1.00	1.50	0.70	0.068

Hielo (NHL, por sus siglas en inglés) de la *Hockey Encyclopedia*, las cuales se muestran en la tabla que aparece abajo.

Las variables y las definiciones son las siguientes:

Ganados — Número de juegos ganados en una temporada.

Puntos — Puntos conseguidos en una temporada. Dos puntos por ganar un juego, un punto por perder en tiempo extra, cero puntos por perder en tiempo regular.

GF — Tantos a favor. Total de tantos anotados durante la temporada.

GA — Tantos en contra. Tantos marcados en contra del equipo durante la temporada.

PPG — Tantos en jugadas de poder. Puntos anotados en una jugada de poder.

PPcT — Porcentaje de jugadas de poder. El número de tantos en jugadas de poder dividido por el número de oportunidades de jugadas de poder.

SHG — Tantos anotados en inferioridad numérica durante la temporada.

PPGA — Tantos en contra en jugadas de poder.

PKPcT — Porcentaje para contrarrestar castigos. Mide la habilidad de un equipo para no permitir tantos cuando su oponente está en una jugada de poder. Tantos de los oponentes en jugadas de poder divididos por las oportunidades de los oponentes.

SHGA — Tantos recibidos en inferioridad numérica.

Ajuste un modelo de regresión lineal múltiple que relacione los juegos ganados con las demás variables.

11-12. Considérese el modelo de regresión lineal

$$Y_i = \beta_0' + \beta_1 (x_{i1} - \bar{x}_1) + \beta_2 (x_{i2} - \bar{x}_2) + \varepsilon_i$$

donde $\bar{x}_1 = \Sigma x_{i1} / n$ y $\bar{x}_2 = \Sigma x_{i2} / n$.

a) Escriba las ecuaciones normales de mínimos cuadrados para este modelo.

b) Verifique que la estimación de mínimos cuadrados de la ordenada al origen de este modelo es $\hat{\beta}_0' = \Sigma y_i / n = \bar{y}$.

c) Suponga que se usa $y_i - \bar{y}$ como variable de respuesta en el modelo anterior. ¿Qué efecto tendrá esto sobre la estimación de mínimos cuadrados de la ordenada al origen?

Equipo	Ganados	Puntos	GF	GA	PPG	PPcT	SHG	PPGA	PkPcT	SHGA
Chicago	47	104	338	268	86	27.2	4	71	76.6	6
Minnesota	40	96	321	290	91	26.4	17	67	80.7	20
Toronto	28	68	23	330	79	22.3	13	83	75	9
San Luis	25	65	285	316	67	21.2	9	63	81.3	12
Detroit	21	57	263	344	37	19.3	7	80	72.6	9
Edmonton	47	106	424	315	86	29.3	22	89	77.5	6
Calgary	32	78	321	317	90	27	7	59	77.1	6
Vancouver	30	75	303	309	65	23.8	5	56	80.8	13
Winnipeg	33	74	311	333	78	23.6	10	67	72.8	7
Los Ángeles	27	66	308	365	81	23.8	10	94	68.2	14
Filadelfia	49	106	326	240	60	21.6	15	61	82	7
NY Islanders	42	96	302	226	69	25.8	10	55	83.4	3
Washington	39	94	306	283	75	20.9	3	53	81.6	11
NY Rangers	35	80	306	287	71	22.4	12	75	76	8
Nueva Jersey	17	48	230	338	66	21.9	74	78	73.5	10
Pittsburgh	18	45	257	394	81	22.6	3	110	72.2	15
Boston	50	110	327	228	67	22.2	8	53	80.7	6
Montreal	42	98	350	286	64	22.2	8	68	73.8	8
Búffalo	38	89	318	285	67	21.5	12	48	82.5	9
Quebec	34	80	343	336	61	20.7	6	92	73.6	6
Hartford	19	45	261	403	51	19.3	6	70	76.1	9

11-4 PROPIEDADES DE LOS ESTIMADORES DE MÍNIMOS CUADRADOS Y ESTIMACIÓN DE σ^2

Las propiedades estadísticas de los estimadores de mínimos cuadrados $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ pueden encontrarse fácilmente, bajo ciertos supuestos sobre los términos del error ε_1, ε_2, ..., ε_n, del modelo de regresión. Paralelamente a los supuestos hechos en el capítulo 10, se supone que los errores ε_i son estadísticamente independientes con media cero y varianza σ^2. Bajo estos supuestos, los estimadores de mínimos cuadrados $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ son estimadores *insesgados* de los coeficientes de regresión β_0, β_1, ..., β_k. Esta propiedad puede establecerse como sigue:

$$
\begin{aligned}
E(\hat{\boldsymbol{\beta}}) &= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\right] \\
&= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})\right] \\
&= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right] \\
&= \boldsymbol{\beta}
\end{aligned}
$$

puesto que $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ y $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I}$, la matriz identidad. Por tanto, $\hat{\boldsymbol{\beta}}$ es un estimador insesgado de $\boldsymbol{\beta}$.

Las varianzas de las $\hat{\boldsymbol{\beta}}$ se expresan en términos de los elementos de la inversa de la matriz $\mathbf{X}'\mathbf{X}$. La inversa de $\mathbf{X}'\mathbf{X}$ multiplicada por la constante σ^2 representa la **matriz de covarianza** de los coeficientes de regresión $\hat{\boldsymbol{\beta}}$. Los elementos de la diagonal de $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ son las varianzas de $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$, y los elementos que están fuera de la diagonal de esta matriz son las covarianzas. Por ejemplo, si se tienen $k = 2$ variables regresoras, como en el problema de la resistencia al desprendimiento, entonces

$$
\begin{aligned}
\mathbf{C} &= (\mathbf{X}'\mathbf{X})^{-1} \\
&= \begin{bmatrix} C_{00} & C_{01} & C_{02} \\ C_{10} & C_{11} & C_{12} \\ C_{20} & C_{21} & C_{22} \end{bmatrix}
\end{aligned}
$$

que es simétrica ($C_{10} = C_{01}$, $C_{20} = C_{02}$, y $C_{21} = C_{12}$) porque $(\mathbf{X}'\mathbf{X})^{-1}$ es simétrica, y se tiene

$$
\begin{aligned}
V(\hat{\beta}_j) &= \sigma^2 C_{jj}, \quad j = 0, 1, 2 \\
\operatorname{cov}(\hat{\beta}_i, \hat{\beta}_j) &= \sigma^2 C_{ij}, \quad i \neq j
\end{aligned}
$$

En general, la matriz de covarianza de $\hat{\boldsymbol{\beta}}$ es una matriz simétrica ($p \times p$) cuyo elemento *jj*-ésimo es la varianza de $\hat{\beta}_j$, y cuyo elemento *i*, *j*-ésimo es la covarianza entre $\hat{\beta}_i$ y $\hat{\beta}_j$, es decir,

$$
\operatorname{cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2\mathbf{C}
$$

Las estimaciones de las varianzas de estos coeficientes de regresión se obtienen sustituyendo σ^2 con una estimación apropiada. Cuando σ^2 se sustituye con una estimación $\hat{\sigma}^2$, a la raíz cuadra-

da de la varianza estimada del coeficiente de regresión j-ésima se le llama el **error estándar estimado** de $\hat{\beta}_j$, o $se(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 C_{jj}}$. Como en el caso de una regresión lineal simple, la estimación de σ^2 se define en términos de la suma de los cuadrados de los residuales

$$SS_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$$

Puesto que puede demostrarse que el valor esperado de la cantidad SS_E es $\sigma^2(n-p)$, un estimador insesgado de σ^2 está dado por el **cuadrado medio de error** o **residual**

$$\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-p} \tag{11-16}$$

El cuadrado medio de error, MS_E, aparecerá en la salida impresa de la mayoría de los programas de computadora de regresión múltiple. Sin embargo, es sencillo obtener una fórmula para calcularlo. Puede escribirse

$$SS_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2 = \mathbf{e}'\mathbf{e}$$

Al sustituir $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, puede demostrarse que

$$SS_E = \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} \tag{11-17}$$

Por lo tanto, otra forma de escribir la ecuación 11-16 es

$$\hat{\sigma}^2 = \frac{SS_E}{n-p} = \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-p}$$

EJEMPLO 11-3

Se estimará la varianza del error σ^2 para el problema de regresión múltiple de los ejemplos 11-1 y 11-2. Usando los datos de la variable de respuesta (la resistencia al desprendimiento) de la tabla 11-2, se encuentra

$$\mathbf{y}'\mathbf{y} = \sum_{i=1}^{25} y_i^2 = 27\,177.9510$$

y

$$\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} = \begin{bmatrix} 2.26379143 & 2.744226964 & 0.01252781 \end{bmatrix} \begin{bmatrix} 725.82 \\ 8\,008.37 \\ 274\,811.31 \end{bmatrix}$$

$$= 27\,062.7775$$

Por lo tanto, la suma de los cuadrados del error se encuentra utilizando la ecuación 11-17

$$\begin{aligned} SS_E &= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}\mathbf{X}'\mathbf{y} \\ &= 27\,177.9510 - 27\,062.7775 = 115.1735 \end{aligned}$$

La estimación de σ^2 se calcula con la ecuación 11-16 como

$$\hat{\sigma}^2 = \frac{SS_E}{n-p} = \frac{115.1735}{25-3} = 5.2352$$

11-5 PRUEBAS DE HIPÓTESIS EN LA REGRESIÓN LINEAL MÚLTIPLE

En los problemas de regresión lineal múltiple, ciertas pruebas de hipótesis acerca de los parámetros del modelo son útiles para medir la adecuación del modelo. En esta sección se describen varios procedimientos importantes para probar hipótesis. Como en el caso de la regresión lineal simple, la prueba de hipótesis requiere que los términos del error ε_i del modelo de regresión tengan una distribución normal e independiente con media cero y varianza σ^2.

11-5.1 Prueba para la significación de una regresión

La prueba de la significación de una regresión sirve para determinar si existe una relación lineal entre la variable de respuesta y y un subconjunto de las variables regresoras $x_1, x_2, \ldots, x_k$. Las hipótesis apropiadas son

$$\begin{aligned} &H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \\ &H_1 : \beta_j \neq 0 \quad \text{para al menos una } j \end{aligned} \tag{11-18}$$

El rechazo de H_0: $\beta_1 = \beta_2 = \cdots = \beta_k = 0$ implica que al menos una de las variables regresoras x_1, $x_2, \ldots, x_k$ contribuye de manera significativa al modelo.

La prueba para la significación de una regresión es una generalización del procedimiento usado en una regresión lineal simple. Se hace la partición de la suma total de los cuadrados SS_T, en una suma de los cuadrados debida a la regresión y una suma de cuadrados debida al error, es decir,

$$SS_T = SS_R + SS_E$$

y si H_0: $\beta_1 = \beta_2 = \cdots = \beta_k = 0$ es verdadera, entonces SS_R/σ^2 es una variable aleatoria ji-cuadrada con k grados de libertad. Obsérvese que el número de grados de libertad de esta variable aleatoria ji-cuadrada es igual al número de variables regresoras del modelo. También puede demostrarse

que SS_E/σ^2 es una variable aleatoria ji-cuadrada con $n-p$ grados de libertad, y que SS_E y SS_R son independientes. El estadístico de la prueba para H_0: $\beta_1 = \beta_2 = \cdots = \beta_k = 0$ es

$$F_0 = \frac{SS_R/k}{SS_E/(n-p)} = \frac{MS_R}{MS_E} \tag{11-19}$$

Deberá rechazarse H_0 si el valor calculado del estadístico de prueba de la ecuación 11-19, f_0, es mayor que $f_{\alpha, k, n-p}$. Se acostumbra resumir el procedimiento en una tabla de análisis de varianza como la tabla 11-5.

En la ecuación 11-17 se ha deducido una fórmula para calcular SS_E; es decir,

$$SS_E = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$$

Es sencillo encontrar una fórmula para calcular SS_R. Puesto que $SS_T = \Sigma_{i=1}^n y_i^2 - \left(\Sigma_{i=1}^n y_i\right)^2/n = \mathbf{y}'\mathbf{y} - \left(\Sigma_{i=1}^n y_i\right)^2/n$ la ecuación anterior puede escribirse como

$$SS_E = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} - \left[\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}\right]$$

o bien

$$SS_E = SS_T - SS_R$$

Por lo tanto, la suma de cuadrados de la regresión es

$$SS_R = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \tag{11-20}$$

Tabla 11-5 Análisis de varianza para probar la significación de la regresión en una regresión múltiple

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Regresión	SS_R	k	MS_R	MS_R/MS_E
Error o residual	SS_E	$n-p$	MS_E	
Total	SS_T	$n-1$		

EJEMPLO 11-4

Se probará la significación de la regresión (con $\alpha = 0.05$) utilizando los datos de la resistencia al desprendimiento del alambre adherido del ejemplo 11-1. Algunas de las cantidades numéricas que se necesitan se calcularon en el ejemplo 11-3. La suma total de cuadrados es

$$SS_T = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$$

$$= 27\,177.9510 - \frac{(725.82)^2}{25} = 6\,105.9447$$

La suma de cuadrados de la regresión se calcula con la ecuación 11-20 como sigue:

$$SS_R = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$$

$$= 27\,062.7775 - \frac{(725.82)^2}{25} = 5\,990.7712$$

y por sustracción

$$SS_R = SS_T - SS_R$$

$$= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} = 115.1735$$

El análisis de varianza se muestra en la tabla 11-6. Para probar H_0: $\beta_1 = \beta_2 = 0$, se calcula el estadístico

$$f_0 = \frac{MS_R}{MS_E} = \frac{2\,995.3856}{5.2352} = 572.17$$

Puesto que $f_0 > f_{0.05,\,2,\,22} = 3.44$ (o puesto que el valor P es considerablemente menor que $\alpha = 0.05$), se rechaza la hipótesis nula y se concluye que la resistencia al desprendimiento se relaciona linealmente ya sea con la longitud del alambre, con la altura de la matriz, o con ambas. Sin embargo, cabe señalar que esto no implica necesariamente que la relación encontrada es un modelo apropiado para predecir la resistencia al desprendimiento como una función de la longitud del alambre y la altura de la matriz. Es necesario realizar pruebas de adecuación adicionales antes de que el uso de este modelo pueda considerarse apropiado en la práctica.

Tabla 11-6 Prueba de la significación de la regresión del ejemplo 11-4

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Regresión	5 990.7712	2	2 995.3856	572.17	1.08E-19
Error o residual	115.1735	22	5.2352		
Total	6 105.9447	24			

La mayoría de los programas de computadora de regresión múltiple incluyen la prueba de la significación de la regresión en la salida impresa. La parte inferior de la tabla 11-4 es la salida del programa Minitab para este ejemplo. Compárense las tablas 11-4 y 11-6 y obsérvese la equivalencia. El valor P se redondea a cero en la salida de la computadora.

11-5.2 Pruebas para los coeficientes de regresión individuales y para subconjuntos de coeficientes

Con frecuencia existe interés en probar hipótesis sobre los coeficientes de regresión individuales. Estas pruebas serían útiles para determinar el valor potencial de cada una de las variables regresoras del modelo de regresión. Por ejemplo, tal vez el modelo sería más efectivo con la inclusión de variables adicionales, o quizá con la eliminación de una o más de las variables regresoras presentes en el modelo.

La incorporación de una variable en un modelo de regresión siempre hace que la suma de cuadrados de la regresión se incremente y que la suma de cuadrados del error se decremente. Es necesario decidir si el incremento de la suma de cuadrados de la regresión es lo suficientemente grande para justificar el uso de una variable adicional en el modelo. Además, la incorporación de una variable de escasa importancia al modelo en realidad puede incrementar el cuadrado medio de error, lo cual indicaría que la inclusión de dicha variable en realidad demeritó el ajuste del modelo a los datos.

Las hipótesis para probar la significación de cualquier coeficiente de regresión individual, por ejemplo β_j, son

$$H_0 = \beta_j = 0$$
$$H_1 = \beta_j \neq 0 \qquad (11\text{-}21)$$

Si no se rechaza H_0: $\beta_j = 0$, entonces esto indica que la variable regresora x_j puede eliminarse del modelo. El estadístico de la prueba para esta hipótesis es

$$T_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \qquad (11\text{-}22)$$

donde C_{jj} es el elemento de la diagonal de $(\mathbf{X'X})^{-1}$ que corresponde a $\hat{\beta}_j$. Obsérvese que el denominador de la ecuación 11-22 es el error estándar del coeficiente de regresión $\hat{\beta}_j$. La hipótesis nula H_0: $\beta_j = 0$ se rechaza si $|t_0| > t_{\alpha/2,\, n-p}$. A ésta se le llama **prueba parcial** o **marginal**, debido a que el coeficiente de regresión $\hat{\beta}_j$ depende de todas las demás variables regresoras x_i $(i \neq j)$ que están en el modelo. Se ampliará el tema en el siguiente ejemplo.

EJEMPLO 11-5

Considérense los datos de la resistencia al desprendimiento del alambre adherido, y suponga que quiere probarse la hipótesis de que el coeficiente de regresión de x_2 (la altura de la matriz) es cero. Las hipótesis son

$$H_0 : \beta_2 = 0$$
$$H_1 : \beta_2 \neq 0$$

El elemento de la diagonal principal de la matriz $(\mathbf{X'X})^{-1}$ que corresponde a $\hat{\beta}_2$ es C_{22} = 0.0000015, por lo que el estadístico t de la ecuación 11-24 es

$$t_0 = \frac{\hat{\beta}_2}{\sqrt{\hat{\sigma}^2 C_{22}}} = \frac{0.01253}{\sqrt{(5.2352)(0.0000015)}} = 4.4767$$

Puesto que $t_{0.025,\,22} = 2.074$, se rechaza H_0: $\beta_2 = 0$ y se concluye que la variable x_2 (la altura de la matriz) contribuye de manera significativa al modelo. También pudo haberse usado un valor P para llegar a una conclusión. El valor P para $t_0 = 4.4767$ es $P = 0.0002$, por lo que con $\alpha = 0.05$ se habría rechazado la hipótesis nula. Obsérvese que esta prueba mide la contribución parcial o marginal de x_2 dado que x_1 está en el modelo. Es decir, la prueba t mide la contribución de agregar la variable x_2 = altura de la matriz a un modelo que contiene ya a x_1 = longitud del alambre. En la tabla 11-4 se muestra el valor de la prueba t calculado con el programa Minitab. Obsérvese que la computadora produce una prueba t para cada coeficiente de regresión del modelo. Estas pruebas t indican que ambas variables regresoras contribuyen al modelo.

Hay otra manera de probar la contribución de una variable regresora individual al modelo. En este enfoque se determina el incremento en la suma de los cuadrados de regresión que se obtiene al agregar una variable x_j (por ejemplo) al modelo, dado que las otras x_i $(i \neq j)$ variables están incluidas ya en la ecuación de regresión.

Al procedimiento usado para este fin se le llama la **prueba general de la significación de la regresión**, o el método de **suma extra de los cuadrados**. Este procedimiento también puede usarse para investigar la contribución de un *subconjunto* de las variables regresoras al modelo. Considérese el modelo de regresión con k variables regresoras

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{11-23}$$

donde $\mathbf{y}$ es $(n \times 1)$, $\mathbf{X}$ es $(n \times p)$, $\boldsymbol{\beta}$ es $(p \times 1)$, $\boldsymbol{\varepsilon}$ es $(n \times 1)$ y $p = k + 1$. Sería conveniente determinar si el subconjunto de las variables regresoras $x_1, x_2, \ldots, x_r$ $(r < k)$ en conjunto, contribuye de manera significativa al modelo de regresión. Sea que se hace la partición del vector de los coeficientes de regresión de la siguiente manera:

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} \tag{11-24}$$

donde $\boldsymbol{\beta}_1$ es $(r \times 1)$ y $\boldsymbol{\beta}_2$ es $[(p - r) \times 1)]$. Quiere probarse las hipótesis

$$H_0 : \boldsymbol{\beta}_1 = \mathbf{0}$$
$$H_1 : \boldsymbol{\beta}_1 \neq \mathbf{0} \tag{11-25}$$

donde $\mathbf{0}$ denota un vector de ceros. El modelo puede escribirse como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \tag{11-26}$$

donde $\mathbf{X}_1$ representa las columnas de $\mathbf{X}$ asociadas con $\boldsymbol{\beta}_1$ y $\mathbf{X}_2$ representa las columnas de $\mathbf{X}$ asociadas con $\boldsymbol{\beta}_2$.

Para el **modelo completo** (incluyendo tanto a $\boldsymbol{\beta}_1$ como a $\boldsymbol{\beta}_2$), se sabe que $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Además, la suma de los cuadrados de la regresión para todas la variables, incluyendo la ordenada al origen, es

$$SS_R(\boldsymbol{\beta}) = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \qquad (p = k+1 \text{ grados de libertad})$$

y

$$MS_E = \frac{\mathbf{y}'\mathbf{y} - \boldsymbol{\beta}\mathbf{X}'\mathbf{y}}{n-p}$$

A $SS_R(\boldsymbol{\beta})$ se le llama la suma de los cuadrados de la regresión debida a $\boldsymbol{\beta}$. Para encontrar la contribución de los términos en $\boldsymbol{\beta}_1$ a la regresión, se ajusta el modelo suponiendo que es verdadera la hipótesis nula H_0: $\boldsymbol{\beta}_1 = \mathbf{0}$. El **modelo reducido** se encuentra por la ecuación 11-26 como

$$\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \tag{11-27}$$

La estimación de mínimos cuadrados de $\boldsymbol{\beta}_2$ es $\hat{\boldsymbol{\beta}}_2 = (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{y}$, y

$$SS_R(\boldsymbol{\beta}_2) = \hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\,\mathbf{y} \quad (p-r \text{ grados de libertad}) \tag{11-28}$$

La suma de cuadrados de la regresión debida a $\boldsymbol{\beta}_1$ dado que $\boldsymbol{\beta}_2$ está ya en el modelo es

$$SS_R(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2) = SS_R(\boldsymbol{\beta}) - SS_R(\boldsymbol{\beta}_2) \tag{11-29}$$

Esta suma de los cuadrados tiene r grados de libertad. En ocasiones se le llama la suma extra de los cuadrados debida a $\boldsymbol{\beta}_1$. Obsérvese que $SS_R(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2)$ es el incremento en la suma de los cuadrados de la regresión debida a la inclusión de las variables $x_1, x_2, \ldots, x_r$ en el modelo. Entonces $SS_R(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2)$ es independiente de MS_E, y la hipótesis nula $\boldsymbol{\beta}_1 = \mathbf{0}$ puede probarse con el estadístico

$$F_0 = \frac{SS_R(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2)/r}{MS_E} \tag{11-30}$$

Si el valor calculado del estadístico de la prueba $f_0 > f_{\alpha, r, n-p}$, se rechaza H_0, y se concluye que al menos uno de los parámetros en $\boldsymbol{\beta}_1$ es diferente de cero y, por consiguiente, al menos una de las variables $x_1, x_2, \ldots, x_r$ en $\mathbf{X}_1$ contribuye de manera significativa al modelo de regresión. Algunos autores llaman a la prueba de la ecuación 11-30 **prueba *F* parcial**.

La prueba *F* parcial es muy útil. Puede usarse para medir la contribución de cada variable regresora individual x_j como si se tratara de la última variable agregada al modelo, calculando

$$SSR(\beta_j \mid \beta_0, \beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k), \quad j = 1, 2, \ldots, k$$

Éste es el incremento en la suma de cuadrados de la regresión debido a la incorporación de x_j a un modelo que incluya ya a $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k$. La prueba *F* parcial es un procedimiento más general por cuanto puede medir el efecto de conjuntos de variables. En la sección 11-10 se explica la forma en que la prueba *F* parcial desempeña un papel importante en la *construcción de modelos*; es decir, en la búsqueda del mejor conjunto de variables regresoras que pueden usarse en el modelo.

EJEMPLO 11-6

Considérense los datos de la resistencia al desprendimiento del alambre adherido del ejemplo 11-1. Se investigará la contribución de la variable x_2 (la altura de la matriz) al modelo usando el enfoque de la prueba *F* parcial. Es decir, quiere probarse

$$H_0 : \beta_2 = 0$$
$$H_i : \beta_2 \neq 0$$

Para probar esta hipótesis, se necesita la suma extra de los cuadrados debida a β_2, o

$$\begin{aligned} SS_R(\beta_2 \mid \beta_1, \beta_0) &= SS_R(\beta_1, \beta_2, \beta_0) - SS_R(\beta_1, \beta_0) \\ &= SS_R(\beta_1, \beta_2 \mid \beta_0) - SS_R(\beta_1 \mid \beta_0) \end{aligned}$$

En el ejemplo 11-4 se calculó

$$SS_R(\beta_1, \beta_2 \beta_0) = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)}{n} = 5990.7712 \quad \text{(2 grados de libertad)}$$

y por el ejemplo 10-10, donde se ajustó el modelo $Y = \beta_0 + \beta_1 x_1 + \varepsilon$, puede calcularse

$$\begin{aligned} SS_R(\beta_1 \mid \beta_0) &= \hat{\beta}_1 S_{xy} = (2.9027)(2\,027.7132) \\ &= 5\,885.8521 \quad \text{(1 grado de libertad)} \end{aligned}$$

Por lo tanto,

$$\begin{aligned} SS_R(\beta_2 \mid \beta_1, \beta_0) &= 5\,990.7712 - 5\,885.8521 \\ &= 104.9191 \quad \text{(1 grado de libertad)} \end{aligned}$$

Éste es el incremento en la suma de cuadrados de la regresión debido a la inclusión de x_2 en un modelo que contenía ya a x_1. Para probar H_0: $\beta_2 = 0$, se calcula el estadístico de la prueba

$$f_0 = \frac{SS_R(\beta_2 \mid \beta_1, \beta_0)/1}{MS_E} = \frac{104.9191/1}{5.2352} = 20.04$$

Obsérvese que MS_E del modelo completo, utilizando tanto x_1 como x_2, se usa en el denominador del estadístico de prueba. Puesto que $f_{0.05,1,22} = 4.30$, se rechaza H_0: $\beta_2 = 0$, y se concluye que la variable regresora de la altura de la matriz (x_2) contribuye de manera significativa al modelo.

En la tabla 11-4 se muestra la salida del programa Minitab para la regresión de los datos de la resistencia al desprendimiento del alambre adherido. En la parte inferior de esta tabla, la cantidad rotulada "Seq SS" muestra la suma de los cuadrados obtenida al ajustar sólo x_1 (5 885.9) y la suma de los cuadrados obtenida al ajustar x_2 después de x_1. En cuanto a la notación empleada, aquí se utilizó $SS_R(\beta_1 \mid \beta_0)$ y $SS_R(\beta_2 \mid \beta_1, \beta_0)$ para referirse a las anteriores cantidades.

Puesto que la prueba F parcial del ejemplo anterior incluye una sola variable, es equivalente a la prueba t. Para ver esto, recuérdese por el ejemplo 11-6 que la prueba t de H_0: $\beta_2 = 0$ resultó en el estadístico de prueba $t_0 = 4.4767$. Además, el cuadrado de una variable aleatoria t con ν grados de libertad es una variable aleatoria F con uno y ν grados de libertad, y se observa que $t_0^2 = (4.4767)^2 = 20.04 = f_0$.

EJERCICIOS DE LAS SECCIONES 11-4 Y 11-5

11-13. Considérese el ajuste del modelo de regresión a los datos de la resistencia al corte del terreno del ejercicio 11-1.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

b) Estime σ^2.

c) Estime los errores estándar de los coeficientes de regresión $\hat{\beta}_1$ y $\hat{\beta}_2$.

d) Construya la prueba t para cada coeficiente de regresión. ¿Cuáles son las conclusiones, usando $\alpha = 0.05$?

11-14. Considérense los datos del índice de absorción del ejercicio 11-2. La suma total de los cuadrados para y es $S_{yy} = 742.00$.

a) Pruebe la significación de la regresión usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?

b) Estime σ^2.

c) Estime el error estándar de $\hat{\beta}_1$.

d) Pruebe la hipótesis H_0: $\beta_1 = 0$ contra H_1: $\beta_1 \neq 0$ usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba? ¿Qué conclusiones pueden sacarse acerca de la utilidad de x_1 como variable regresora en este modelo?

11-15. Considérense los datos de la NFL del ejercicio 11-4.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

b) Estime σ^2.

c) Estime el error estándar de $\hat{\beta}_2$.

d) Realice la prueba t para los coeficientes de regresión β_2, β_7 y β_8. Usando $\alpha = 0.05$, ¿qué conclusiones pueden sacarse sobre las variables de este modelo?

11-16. Considérense de nuevo los datos de la NFL del ejercicio 11-4.

a) Encuentre la cantidad en que la variable regresora x_8 (yardas por carrera de los

oponentes) incrementa la suma de cuadrados de la regresión.

b) Use los resultados del inciso a) anterior y el ejercicio 11-15 para realizar una prueba F para H_0: $\beta_8 = 0$ contra H_1: $\beta_8 \neq 0$ usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba? ¿Qué conclusiones pueden sacarse?

11-17. Considérense los datos del rendimiento de gasolina del ejercicio 11-5.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?

b) Encuentre el estadístico t de la prueba para ambas variables regresoras. Usando $\alpha = 0.05$, ¿qué conclusiones pueden sacarse? ¿Ambas variables regresoras contribuyen al modelo?

11-18. Se ha ajustado un modelo de regresión $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$ a una muestra de $n = 25$ observaciones. Los cocientes t calculados $\hat{\beta}_j / se(\hat{\beta}_j)$, $j = 1,2,3$, son los siguientes: para β_1, $t_0 = 4.82$, para β_2, $t_0 = 8.21$, y para β_3, $t_0 = 0.98$.

a) Encuentre el valor P de cada uno de los estadísticos t.

b) Usando $\alpha = 0.05$, ¿qué conclusiones pueden sacarse acerca de la variable regresora x_3? ¿Parece probable que esta variable regresora contribuya de manera significativa al modelo?

11-19. Considérense los datos del consumo de energía eléctrica del ejercicio 11-6.

a) Pruebe la significación de la regresión usando $\alpha = 0.01$. ¿Cuál es el valor P para esta prueba?

b) Estime σ^2.

c) Use la prueba t para determinar la contribución de cada variable regresora al modelo. Usando $\alpha = 0.01$, ¿qué conclusiones pueden sacarse?

11-20. Considérense los datos del desgaste del rodamiento del ejercicio 11-7 sin interacción.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba? ¿A qué conclusiones se llega?

b) Calcule el estadístico t de cada coeficiente de regresión. Usando $\alpha = 0.05$, ¿qué conclusiones pueden sacarse?

c) Use el método de la suma extra de los cuadrados para investigar la verdadera utilidad de agregar x_2 = carga, a un modelo que contiene ya a x_1 = viscosidad del aceite. Use $\alpha = 0.05$.

11-21. Considérense de nuevo los datos del desgaste del rodamiento de los ejercicios 11-7 y 11-20.

a) Ajuste nuevamente el modelo con un término de interacción. Pruebe la significación de la regresión usando $\alpha = 0.05$.

b) Use el método de la suma extra de los cuadrados para determinar si el término de interacción contribuye de manera significativa al modelo. Use $\alpha = 0.05$.

c) Estime σ^2 del modelo con interacción. Compare el resultado con la estimación de σ^2 en el modelo del ejercicio 11-20.

11-22. Considérense los datos de la resistencia al desprendimiento del alambre adherido del ejercicio 11-8.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. Encuentre el valor P para esta prueba. ¿Qué conclusiones pueden sacarse?

b) Estime σ^2 para este modelo.

c) Calcule el estadístico de la prueba t para cada coeficiente de regresión. Usando $\alpha = 0.05$, ¿qué conclusiones pueden sacarse? ¿Todas las variables contribuyen al modelo?

11-23. Considérense de nuevo los datos del semiconductor del ejercicio 11-9.

a) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?

b) Estime σ^2 para este modelo.

c) Calcule el estadístico de la prueba t para cada coeficiente de regresión. Usando $\alpha = 0.05$, ¿qué conclusiones pueden sacarse?

11-24. En el ejercicio 11-10 se presentan los datos del tratamiento térmico de engranes.

a) Pruebe el modelo de regresión para la significación de la regresión. Usando $\alpha = 0.05$, encuentre el valor P para la prueba y derive conclusiones.
b) Estime σ^2 para este modelo.
c) Evalúe la contribución de cada variable regresora al modelo usando la prueba t con $\alpha = 0.05$.
d) Ajuste un nuevo modelo a la variable de respuesta PITCH usando las nuevas variables regresoras x_1 = SOAKTIME × SOAKPCT y x_2 = DIFFTIME × DIFFPCT.
e) Pruebe el modelo del inciso d) para la significación de la regresión usando $\alpha = 0.05$. Calcule también la prueba t para cada variable regresora y saque conclusiones.
f) Estime σ^2 para el modelo del inciso d) y compare el resultado con la estimación de σ^2 obtenida en el inciso b) anterior. ¿Cuál de las dos estimaciones es menor? ¿Ofrece esto alguna información respecto de cuál de los modelos sería preferible?

11-25. En el ejercicio 11-11 se presentan los datos sobre el desempeño de los equipos de la Liga Nacional de Hockey sobre Hielo.

a) Pruebe el modelo de este ejercicio para la significación de la regresión usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?
b) Use la prueba t para evaluar la contribución de cada variable regresora al modelo. ¿Todas las variables regresoras parecen ser necesarias? Use $\alpha = 0.05$.
c) Ajuste un modelo de regresión que relacione el número de juegos ganados con el número de puntos anotados y con el número de tantos en jugadas de poder. ¿Parece ser ésta una elección lógica de las variables regresoras, considerando la respuesta del inciso b)? Pruebe este nuevo modelo para la significación de la regresión y evalúe la contribución de cada variable regresora al modelo usando la prueba t. Use $\alpha = 0.05$.

11-6 INTERVALOS DE CONFIANZA EN LA REGRESIÓN LINEAL MÚLTIPLE

11-6.1 Intervalos de confianza para los coeficientes individuales de una regresión

En los modelos de regresión múltiple con frecuencia es conveniente construir estimaciones de intervalos de confianza para los coeficientes de regresión $\{\beta_j\}$. El desarrollo de un procedimiento para obtener estos intervalos de confianza requiere que los errores $\{\varepsilon_i\}$ tengan una distribución normal e independiente con media cero y varianza σ^2. Se trata del mismo supuesto requerido en la prueba de hipótesis. Por lo tanto, las observaciones $\{Y_i\}$ tienen una distribución normal e independiente con media $\beta_0 + \sum_{j=1}^{k} \beta_j x_{ij}$ y varianza σ^2. Puesto que el estimador de mínimos cuadrados $\hat{\boldsymbol{\beta}}$ es una combinación lineal de las observaciones, se sigue que $\hat{\boldsymbol{\beta}}$ tiene una distribución normal con vector medio $\boldsymbol{\beta}$ y matriz de covarianza $\sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$. Entonces todos los estadísticos

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \qquad j = 0, 1, \ldots, k \tag{11-31}$$

tienen una distribución t con $n - p$ grados de libertad, donde C_{jj} es el elemento jj-ésimo de la matriz $(\mathbf{X}'\mathbf{X})^{-1}$, y $\hat{\sigma}^2$ es la estimación de la varianza del error, obtenida con la ecuación 11-16. Esto lleva al siguiente intervalo de confianza del 100 $(1 - \alpha)$ por ciento para el coeficiente de regresión β_j, $j = 0, 1, \ldots, k$.

Definición

Un **intervalo de confianza del 100 (1 – α) por ciento para el coeficiente de regresión** β_j, $j = 0, 1, \ldots, k$, en el modelo de regresión lineal múltiple está dado por

$$\hat{\beta}_j - t_{\alpha/2,\, n-p}\sqrt{\hat{\sigma}^2 C_{jj}} \le \beta_j \le \hat{\beta}_j + t_{\alpha/2,\, n-p}\sqrt{\hat{\sigma}^2 C_{jj}} \qquad (11\text{-}32)$$

EJEMPLO 11-7

Se construirá un intervalo de confianza de 95% para el parámetro β_1 del problema de la resistencia al desprendimiento del alambre adherido. La estimación puntual de β_1 es $\hat{\beta}_1 = 2.74427$ y el elemento de la diagonal de $(\mathbf{X'X})^{-1}$ correspondiente a β_1 es $C_{11} = 0.001671$. En el ejemplo 11-3 se obtuvo la estimación 5.2352 para σ^2, y $t_{0.025,\,22} = 2.074$. Por lo tanto, con la ecuación 11-32 el intervalo de confianza de 95% para β_1 se calcula como

$$2.74427 - (2.074)\sqrt{(5.2352)(.001671)} \le \beta_1 \le 2.74427 + (2.074)\sqrt{(5.2352)(.001671)}$$

que se reduce a

$$2.55029 \le \beta_1 \le 2.93825$$

11-6.2 Intervalo de confianza para la respuesta media

También puede obtenerse un intervalo de confianza para la respuesta media en un punto particular, por ejemplo, $x_{01}, x_{02}, \ldots, x_{0k}$. Para estimar la respuesta media en este punto, se define el vector

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \vdots \\ x_{0k} \end{bmatrix}$$

La respuesta media en este punto es $E(Y \mid \mathbf{x}_0) = \mu_{Y|\mathbf{x}_0} = \mathbf{x}_0'\boldsymbol{\beta}$, cuya estimación se hace con

$$\hat{\mu}_{Y|\mathbf{x}_0} = \mathbf{x}_0'\hat{\boldsymbol{\beta}} \qquad (11\text{-}33)$$

Este estimador es insesgado, ya que $E(\mathbf{x}_0'\hat{\boldsymbol{\beta}}) = \mathbf{x}_0'\boldsymbol{\beta} = E(Y|\mathbf{x}_0) = \mu_{Y|\mathbf{x}_0}$ y la varianza de $\hat{\mu}_{Y|\mathbf{x}_0}$ es

$$V(\hat{\mu}_{Y|\mathbf{x}_0}) = \sigma^2\, \mathbf{x}_0'(\mathbf{X'X})^{-1}\mathbf{x}_0 \qquad (11\text{-}34)$$

Un intervalo de confianza del 100(1 – α) por ciento para $\mu_{Y|\mathbf{x}_0}$ puede construirse a partir del estadístico

$$\frac{\hat{\mu}_{Y|\mathbf{x}_0} - \mu_{Y|\mathbf{x}_0}}{\sqrt{\hat{\sigma}^2 \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}} \tag{11-35}$$

Definición

Para el modelo de regresión lineal múltiple, un **intervalo de confianza del 100(1 – α) por ciento para la respuesta media** en el punto $x_{01}, x_{02}, \ldots, x_{0k}$ es

$$\hat{\mu}_{Y|\mathbf{x}_0} - t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}$$

$$\leq \mu_{Y|\mathbf{x}_0} \leq \hat{\mu}_{Y|\mathbf{x}_0} + t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} \tag{11-36}$$

La ecuación 11-36 es un intervalo de confianza alrededor del hiperplano de regresión. Se trata de una generalización de la ecuación 10-33 para una regresión múltiple.

EJEMPLO 11-8

Al ingeniero del ejemplo 11-1 le interesaría construir un intervalo de confianza de 95% para la resistencia media al desprendimiento de un alambre adherido de longitud $x_1 = 8$ y altura de la matriz $x_2 = 275$. Por lo tanto,

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 8 \\ 275 \end{bmatrix}$$

La respuesta media estimada en este punto se encuentra con la ecuación 11-33 como

$$\hat{\mu}_{Y|\mathbf{x}_0} = \mathbf{x}_0'\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1 & 8 & 275 \end{bmatrix} \begin{bmatrix} 2.26379 \\ 2.74427 \\ 0.01253 \end{bmatrix} = 27.66$$

La varianza de $\hat{\mu}_{Y|\mathbf{x}_0}$ se estima con

$$\hat{\sigma}^2 \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0 = 5.2352 \begin{bmatrix} 1 & 8 & 275 \end{bmatrix}$$

$$\times \begin{bmatrix} .214653 & -.007491 & -.000340 \\ -.007491 & .001671 & -.000019 \\ -.000340 & -.000019 & .0000015 \end{bmatrix} \begin{bmatrix} 1 \\ 8 \\ 275 \end{bmatrix}$$

$$= 5.2352\,(0.04444) = 0.23266$$

Por lo tanto, un intervalo de confianza de 95% para la resistencia media al desprendimiento en este punto se encuentra con la ecuación 11-36 como

$$27.66 - 2.074\sqrt{0.23266} \le \mu_{Y|\mathbf{x}_0} \le 27.66 + 2.074\sqrt{0.23266}$$

que se reduce a

$$26.66 \le \mu_{Y|\mathbf{x}_0} \le 28.66$$

11-7 PREDICCIÓN DE NUEVAS OBSERVACIONES

Es posible usar un modelo de regresión para predecir observaciones futuras de la variable de respuesta Y correspondientes a valores particulares de las variables independientes, por ejemplo, $x_{01}, x_{02}, \ldots, x_{0k}$. Si $\mathbf{x}_0' = [1, x_{01}, x_{02}, \ldots, x_{0k}]$, entonces una estimación puntual de la observación futura Y_0 en el punto $x_{01}, x_{02}, \ldots, x_{0k}$ es

$$\hat{y}_0 = \mathbf{x}_0'\hat{\boldsymbol{\beta}} \tag{11-37}$$

Un **intervalo de predicción para esta observación futura** del $100(1 - \alpha)$ por ciento es

$$\hat{y}_0 - t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \le y_0 \le \hat{y}_0 - t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \tag{11-38}$$

Este intervalo de predicción es una generalización del intervalo de predicción dado en la ecuación 10-35 para una observación futura en una regresión lineal simple. Si se compara la ecuación 11-38 del intervalo de predicción con la expresión del intervalo de confianza para la media, la ecuación 11-36, se observará que el intervalo de predicción siempre tiene una anchura mayor que el intervalo de confianza. El intervalo de confianza expresa el error al estimar la media de una distribución, en tanto que el intervalo de predicción expresa el error al predecir una observación futura a partir de la distribución en el punto $\mathbf{x}_0$. Esto debe incluir el error al estimar la media en ese punto, así como la variabilidad inherente de la variable aleatoria Y en el mismo valor $\mathbf{x} = \mathbf{x}_0$.

Cuando se hace la predicción de nuevas observaciones y si se estima la respuesta media en un punto dado $x_{01}, x_{02}, \ldots, x_{0k}$, es necesario tener cuidado al hacer una extrapolación fuera de la región que contiene las observaciones originales. Es altamente probable que un modelo que es un ajuste adecuado en la región de los datos originales deje de serlo fuera de esa región. En una regresión múltiple con frecuencia es fácil hacer una extrapolación inadvertidamente, ya que los

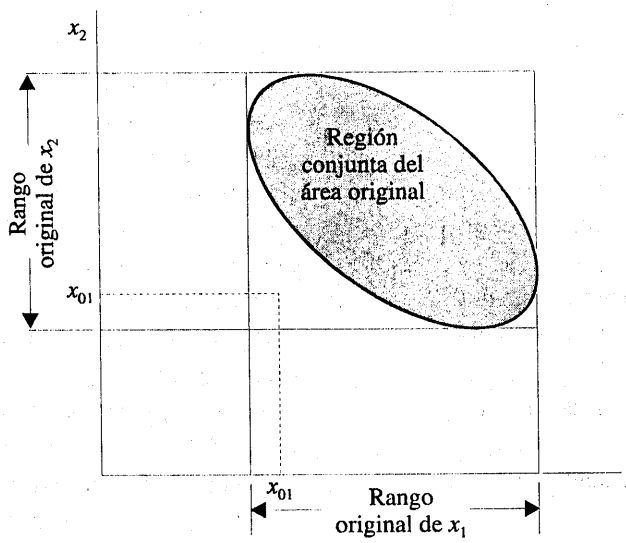


Figura 11-4 Un ejemplo de extrapolación en una regresión múltiple.

niveles de las variables $(x_{i1}, x_{i2}, ..., x_{ik})$, $i = 1, 2, ..., n$, definen conjuntamente la región que contiene los datos. Como un ejemplo, considérese la figura 11-4, donde se ilustra la región que contiene las observaciones de un modelo de regresión de dos variables. Obsérvese que el punto (x_{01}, x_{02}) se localiza dentro de los rangos de las dos variables regresoras x_1 y x_2, pero está fuera de la región que abarcan en realidad las observaciones originales. Por tanto, predecir el valor de una nueva observación o estimar la respuesta media en este punto es una extrapolación del modelo de regresión original.

EJEMPLO 11-9

Suponga que el ingeniero del ejemplo 11-1 desea construir un intervalo de predicción de 95% para la resistencia al desprendimiento del alambre adherido cuando la longitud del alambre es $x_1 = 8$ y la altura de la matriz es $x_2 = 275$. Obsérvese que $\mathbf{x}_0' = [1 \quad 8 \quad 275]$, y que la estimación puntual de la resistencia al desprendimiento es $\hat{y}_0 = \mathbf{x}_0'\hat{\boldsymbol{\beta}} = 27.66$. Además, en el ejemplo 11-8 se calculó $\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0 = 0.04444$. Por lo tanto, por la ecuación 11-38 se tiene

$$27.66 - 2.074\sqrt{5.2352(1+0.04444)} \le y_0 \le 27.66 + 2.074\sqrt{5.2352(1+0.04444)}$$

y el intervalo de predicción de 95% es

$$22.81 \le y_0 \le 32.51$$

Obsérvese que el intervalo de predicción tiene una anchura mayor que el intervalo de confianza para la respuesta media en el mismo punto, el cual se calculó en el ejemplo 11-8.

EJERCICIOS DE LAS SECCIONES 11-6 Y 11-7

11-26. Considérense los datos de la absorción del suelo del ejercicio 11-2.

a) Encuentre un intervalo de confianza de 95% para el coeficiente de regresión β_1.

b) Encuentre un intervalo de confianza de 95% para el índice de absorción medio del suelo cuando $x_1 = 200$ y $x_2 = 50$.

c) Encuentre un intervalo de predicción de 95% para el índice de absorción del suelo cuando $x_1 = 200$ y $x_2 = 50$.

11-27. Considérense los datos de la NFL del ejercicio 11-4.

a) Encuentre un intervalo de confianza de 95% para β_8.

b) ¿Cuál es el error estándar estimado de $\hat{\mu}_{Y|\mathbf{x}_0}$ cuando $x_2 = 2\,000$ yardas, $x_7 = 60\%$ y $x_8 = 1\,800$ yardas?

c) Encuentre un intervalo de confianza de 95% para el número promedio de juegos ganados cuando $x_2 = 2\,000$, $x_7 = 60\%$ y $x_8 = 1\,800$.

11-28. Considérense los datos del rendimiento de gasolina del ejercicio 11-5.

a) Encuentre un intervalo de confianza de 99% para β_1 y β_6.

b) Encuentre un intervalo de confianza de 99% para la media de Y cuando $x_1 = 300$ y $x_6 = 4$.

c) Ajuste un nuevo modelo de regresión a estos datos usando x_1, x_2, x_6 y x_{10} como variables regresoras. Encuentre los intervalos de confianza de 99% para los coeficientes de regresión del nuevo modelo.

d) Compare las longitudes de los intervalos de confianza para β_1 y β_6 del inciso c) con los que se encontraron en el inciso a). ¿Qué intervalos tienen una anchura mayor? ¿Ofrece esto alguna información acerca de haber agregado las variables x_2 y x_{10} al modelo?

11-29. Considérense los datos del consumo de energía eléctrica del ejercicio 11-6.

a) Encuentre los intervalos de confianza de 95% para β_1, β_2, β_3 y β_4.

b) Encuentre un intervalo de confianza de 95% para la media de Y cuando $x_1 = 75$, $x_2 = 24$, $x_3 = 90$ y $x_4 = 98$.

c) Encuentre un intervalo de predicción de 95% para el consumo de energía cuando $x_1 = 75$, $x_2 = 24$, $x_3 = 90$ y $x_4 = 98$.

11-30. Considérense los datos del desgaste del rodamiento del ejercicio 11-7.

a) Encuentre los intervalos de confianza de 99% para β_1 y β_2.

b) Calcule de nuevo los intervalos de confianza del inciso a) después de agregar el término de interacción x_1x_2 al modelo. Compare las longitudes de estos intervalos de confianza con los que se calcularon en el inciso a). ¿Las longitudes de los intervalos proporcionan alguna información acerca de la contribución del término de interacción en el modelo?

11-31. Considérense los datos de la resistencia al desprendimiento del alambre adherido del ejercicio 11-8.

a) Encuentre los intervalos de confianza de 95% para los coeficientes de regresión.

b) Encuentre un intervalo de confianza de 95% para la resistencia media al desprendimiento cuando $x_2 = 20$, $x_3 = 30$, $x_4 = 90$ y $x_5 = 2.0$.

c) Encuentre un intervalo de predicción de 95% para la resistencia al desprendimiento cuando $x_2 = 20$, $x_3 = 30$, $x_4 = 90$ y $x_5 = 2.0$.

11-32. Considérense los datos del semiconductor del ejercicio 11-9.

a) Encuentre los intervalos de confianza de 99% para los coeficientes de regresión.

b) Encuentre un intervalo de predicción de 99% para la HFE cuando $x_1 = 14.5$, $x_2 = 220$ y $x_3 = 5.0$.

c) Encuentre un intervalo de confianza de 99% para la HFE media cuando $x_1 = 14.5$, $x_2 = 220$ y $x_3 = 5.0$.

11-33. Considérense los datos del tratamiento térmico del ejercicio 11-10.

a) Encuentre los intervalos de confianza de 95% para los coeficientes de regresión.

b) Encuentre un intervalo de confianza de 95% para el PITCH medio cuando TEMP = 1650, SOAKTIME = 1.00, SOAKPCT = 1.10, DIFFTIME = 1.00 y DIFFPCT = 0.80.

11-34. Considere de nuevo los datos del tratamiento térmico de los ejercicios 11-10 y 11-24, donde se ajustó un modelo al PITCH usando las variables regresoras x_1 = SOAKTIME × SOAKPCT y x_2 = DIFFTIME × DIFFPCT.

a) Utilizando el modelo con las variables regresoras x_1 y x_2, encuentre un intervalo de confianza de 95% para el PITCH medio cuando SOAKTIME = 1.00, SOAKPCT = 1.10, DIFFTIME = 1.00 y DIFFPCT = 0.80.

b) Compare la longitud de este intervalo de confianza con la longitud del intervalo de confianza para el PITCH medio en el mismo punto del inciso b) del ejercicio 11-33, donde se usó un modelo aditivo en SOAKTIME, SOAKPCT, DIFFTIME y DIFFPCT. ¿Cuál de los intervalos de confianza tiene una longitud menor? ¿Dice esto algo acerca de cuál de los modelos es preferible?

11-35. Considérense los datos de la NHL del ejercicio 11-11.

a) Encuentre un intervalo de confianza de 95% para el coeficiente de regresión de la variable "Puntos".

b) Ajuste un modelo de regresión lineal simple que relacione la variable de respuesta "ganados" con la variable regresora "Puntos".

c) Encuentre un intervalo de confianza de 95% para la pendiente del modelo de regresión lineal simple del inciso b).

d) Compare las longitudes de los dos intervalos de confianza calculados en los incisos a) y c). ¿Cuál de los intervalos de confianza tiene una longitud menor? ¿Dice esto algo acerca de cuál de los modelos es preferible?

11-8 MEDIDAS DE LA ADECUACIÓN DEL MODELO

Es posible usar varias técnicas para medir la adecuación de un modelo de regresión múltiple. En esta sección se presentarán algunas de ellas.

11-8.1 Coeficiente de determinación múltiple (R^2)

El coeficiente de determinación múltiple R^2 se define como

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \tag{11-39}$$

R^2 es una medida de la cantidad en que se reduce la variabilidad del valor de y obtenido al usar las variables regresoras $x_1, x_2, \ldots, x_k$. Como en el caso de la regresión lineal simple, debe tenerse $0 \le R^2 \le 1$. Sin embargo, un valor grande de R^2 no implica necesariamente que el modelo de regresión es adecuado. Agregar una variable al modelo incrementará siempre R^2, independientemente de si la variable adicional es estadísticamente significativa o no. Por tanto, los modelos que tienen valores grandes de R^2 pueden producir predicciones pobres de nuevas observaciones, o estimaciones pobres de la variable de respuesta media.

A la raíz cuadrada positiva de R^2 se le llama el **coeficiente de correlación múltiple** entre y y el conjunto de variables regresoras $x_1, x_2, ..., x_k$. Es decir, R es una medida de la asociación lineal entre y y $x_1, x_2, ..., x_k$. Cuando $k = 1$, esto se reduce a la correlación simple entre y y x.

EJEMPLO 11-10

El coeficiente de determinación múltiple para el ajuste del modelo de regresión a los datos de la resistencia al desprendimiento del alambre adherido del ejemplo 11-1 es

$$R^2 = \frac{SS_R}{SS_T} = \frac{5\,990.7712}{6\,105.9447} = 0.981137$$

Es decir, se explica cerca de 98.11% de la variabilidad en la resistencia al desprendimiento y cuando se usan las dos variables regresoras, la longitud del alambre (x_1) y la altura de la matriz (x_2). En el ejemplo 10-10 se desarrolló un modelo de regresión lineal simple que relaciona y con x_1. El valor de R^2 para este modelo es $R^2 = 0.963954$. Por lo tanto, con la inclusión de la variable x_2 en el modelo, R^2 se ha incrementado de 0.963954 a 0.981137.

11-8.2 Análisis de residuales

Los residuales del modelo de regresión múltiple, definidos por $e_i = y_i - \hat{y}_i$, desempeñan un papel importante para juzgar la adecuación del modelo, justo como ocurre en una regresión lineal simple. Como se señaló en la sección 10-8.1, varias gráficas de los residuales con frecuencia son útiles; éstas se ilustran en el ejemplo 11-11. También es conveniente graficar los residuales contra variables que aún no están en el modelo, pero que son posibles candidatas para incluirlas. Los patrones de estas gráficas, similares a los de la figura 10-7, indican que el modelo puede mejorarse mediante la incorporación de la variable candidata.

EJEMPLO 11-11

En la tabla 11-3 se muestran los residuales para el modelo del ejemplo 11-1, y en la figura 11-5 se presenta la gráfica de probabilidad normal de los residuales. No hay desviaciones marcadas de la normalidad que se perciban de inmediato, aun cuando los residuales más grandes (e_{15} = 5.88 y e_{17} = 4.33) no están muy cerca de la línea recta trazada por los demás residuales.

Los **residuales estandarizados**

$$d_i = \frac{e_i}{\sqrt{MS_E}} = \frac{e_i}{\sqrt{\hat{\sigma}^2}}$$

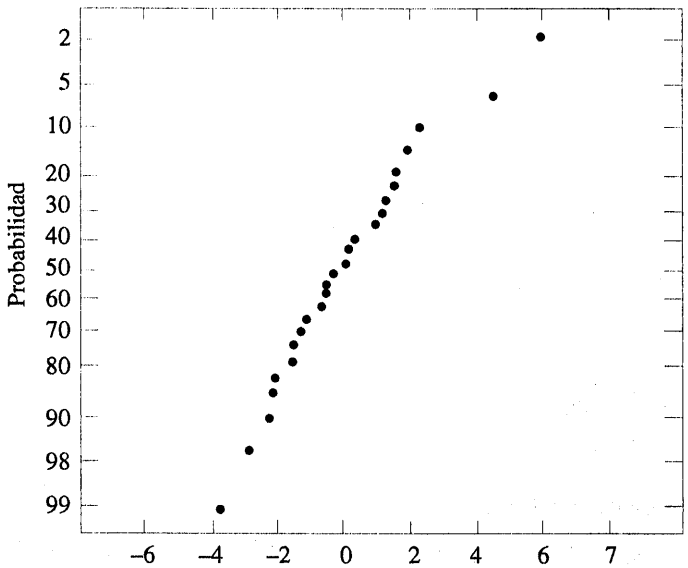


Figura 11-5 Gráfica de probabilidad normal de los residuales.

con frecuencia son más útiles que los residuales ordinarios cuando se determina la magnitud de los residuales. Los residuales estandarizados correspondientes a e_{15} y e_{17} son $d_{15} = 5.88/\sqrt{5.2352} = 2.57$ y $d_{17} = 4.33/\sqrt{4.2352} = 1.89$, y no parecen ser excepcionalmente grandes. La inspección de los datos no revela ningún error al recabar las observaciones 15 y 17, y tampoco produce alguna otra razón para descartar o modificar estos dos puntos.

En la figura 11-6, los residuales se grafican contra $\hat{y}$, y en las figuras 11-7 y 11-8 se grafican contra x_1 y x_2, respectivamente.[1] Los dos residuales más grandes, e_{15} y e_{17}, son evidentes. La figura 11-7 ofrece ciertos indicios de que el modelo subestima la predicción de la resistencia al desprendimiento para las unidades con una longitud del alambre corta ($x_1 \leq 6$) y con una longitud del alambre larga ($x_1 \geq 15$), y que sobrestima la predicción de la resistencia para las unidades con una longitud del alambre intermedia ($7 \leq x_1 \leq 14$). Se obtiene la misma impresión de la figura 11-6. En todo caso, la relación entre la resistencia y la longitud del alambre es lineal (requiriendo la inclusión de un término x_1^2, por ejemplo, en el modelo), o bien requiriendo otras variables regresoras que no se encuentran presentes en el modelo y que afectarían la respuesta. Más adelante se verá que se necesita una tercera variable regresora para modelar adecuadamente estos datos.

En el ejemplo 11-11 se utilizaron los residuales estandarizados $d_i = e_i/\sqrt{\hat{\sigma}^2}$ como una medida de la magnitud de los residuales. Algunos analistas prefieren graficar los residuales estandarizados en lugar de los residuales ordinarios, debido a que los residuales estandarizados se escalan de tal modo que su desviación estándar es aproximadamente la unidad. Por consiguiente, los residuales grandes (que pueden indicar posibles puntos atípicos u observaciones inusuales) serán más evidentes al inspeccionar estas gráficas de los residuales:

[1] Hay otros métodos, descritos en Montgomery y Peck y Myers, en los que se grafica una versión modificada de los residuales, llamada de residual parcial, contra cada variable regresora. Estas gráficas de residual parcial son útiles para mostrar la relación entre la variable de respuesta y y cada variable regresora individual.

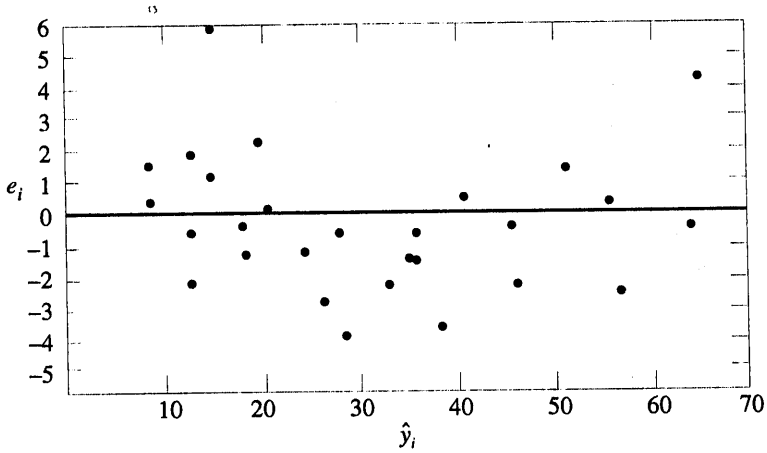


Figura 11-6 Gráfica de los residuales contra $\hat{y}$.

Muchos programas de computadora para regresiones calculan otros tipos de residuales escalados. Uno de los más populares es el **residual studentizado**

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2(1-h_{ii})}} \qquad i=1,2,\ldots,n \tag{11-40}$$

donde h_{ii} es el elemento i-ésimo de la diagonal de la matriz

$$\mathbf{H} = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'$$

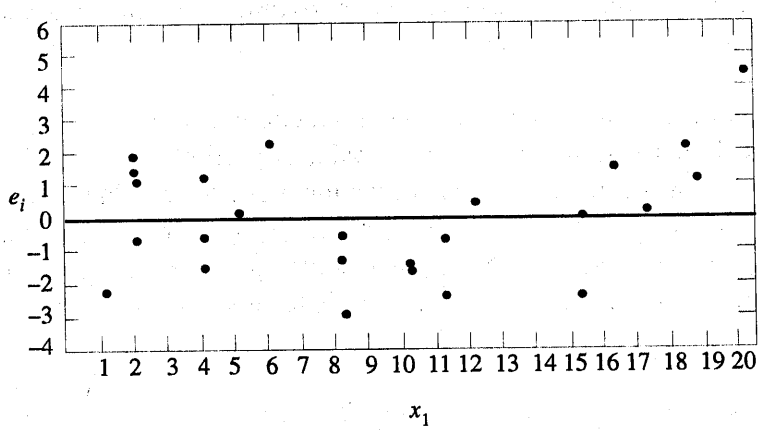


Figura 11-7 Gráfica de los residuales contra x_1.

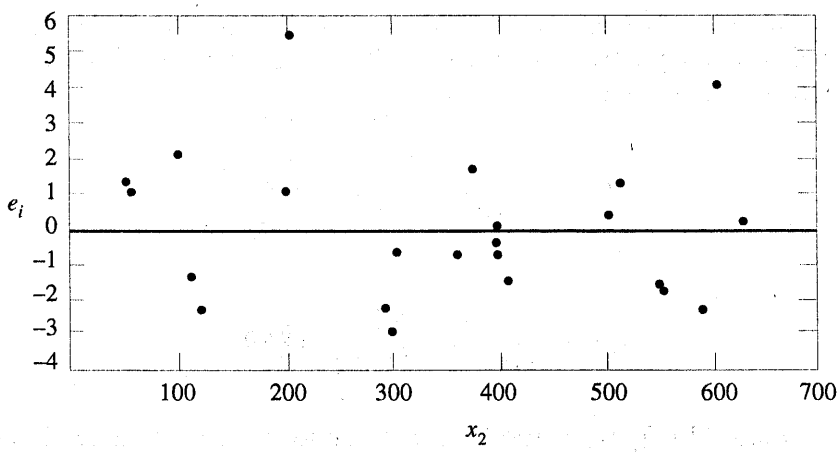


Figura 11-8 Gráfica de los residuales contra x_2.

A la matriz **H** se le llama en ocasiones la matriz "gorro", ya que

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \mathbf{H}\mathbf{y}\end{aligned}$$

Por tanto, **H** transforma los valores observados de **y** en un vector de valores ajustados $\hat{\mathbf{y}}$.

Puesto que cada renglón de la matriz **X** corresponde a un vector, digamos $\mathbf{x}_i' = [1, x_{i1}, x_{i2}, ..., x_{ik}]$, entonces otra manera de escribir los elementos de la diagonal de la matriz gorro es

$$h_{ii} = \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i \tag{11-41}$$

Obsérvese que además de σ^2, h_{ii} es la varianza del valor ajustado $\hat{y}_i$. Las cantidades h_{ii} se usaron en la sección 11-2 para calcular el intervalo de confianza para la respuesta media.

Bajo los supuestos acostumbrados de que los errores del modelo tienen una distribución independiente con media cero y varianza σ^2, puede demostrarse que la varianza del residuo i-ésimo, e_i, es

$$V(e_i) = \sigma^2(1 - h_{ii}), \qquad i = 1, 2, ..., n$$

Además, los elementos h_{ii} deben estar incluidos en el intervalo

$$0 < h_{ii} \leq 1$$

Esto implica que los residuales estandarizados subestiman la verdadera magnitud de los residuales; por tanto, los residuales studentizados serían un estadístico mejor por examinar al evaluar posibles puntos atípicos.

Como ilustración, considérense las dos observaciones del ejemplo 11-11 que se identificaron por tener residuales que podrían ser excepcionalmente grandes, las observaciones 15 y 17. Los residuales estandarizados son

$$d_{15} = \frac{e_{15}}{\sqrt{\hat{\sigma}^2}} = \frac{5.88}{\sqrt{5.2352}} = 2.57$$

y

$$d_{17} = \frac{e_{17}}{\sqrt{MS_E}} = \frac{4.33}{\sqrt{5.2352}} = 1.89$$

Ahora bien $h_{15,\,15} = 0.0737$ y $h_{17,\,17} = 0.2593$, por lo que los residuales studentizados son

$$r_{15} = \frac{e_{15}}{\sqrt{\hat{\sigma}^2\,(1-h_{15,15})}} = \frac{5.88}{\sqrt{5.2352\,(1-0.0737)}} = 2.67$$

y

$$r_{17} = \frac{e_{17}}{\sqrt{\hat{\sigma}^2\,(1-h_{17,\,17})}} = \frac{4.33}{\sqrt{5.2352\,(1-0.2593)}} = 2.20$$

Obsérvese que los residuales studentizados son mayores que los residuales estandarizados correspondientes. Sin embargo, aún así los residuales studentizados no son tan grandes como para despertar preocupaciones serias por la posible presencia de puntos atípicos.

11-8.3 Observaciones influyentes

Cuando se usa una regresión múltiple, algunas veces se encuentra que un subconjunto de las observaciones tiene una influencia fuera de lo común. En ocasiones estas observaciones influyentes se localizan relativamente lejos de la vecindad donde se recabaron el resto de los datos. Una situación hipotética para dos variables se describe en la figura 11-9, donde una de las observaciones del espacio x está muy apartada del resto de los datos. La disposición de los puntos en el espacio x es importante para determinar las propiedades del modelo. Por ejemplo, el punto (x_{i1}, x_{i2}) de la figura 11-9 puede ejercer una gran influencia al determinar R^2, las estimaciones de los coeficientes de regresión o la magnitud del cuadrado medio de error.

Sería conveniente examinar los puntos influyentes a fin de determinar si controlan varias de las propiedades del modelo. Si estos puntos influyentes son puntos "malos" o erróneos de alguna manera, entonces deberán eliminarse. Por otra parte, quizá no haya nada malo con estos puntos, pero al menos sería conveniente determinar si producen o no resultados consistentes con el resto de los datos. De cualquier modo, incluso si un punto influyente es válido, si controla propiedades importantes del modelo, sería conveniente saberlo, ya que podría tener un impacto en el uso del modelo.

Montgomery y Peck y Myers describen varios métodos para detectar observaciones influyentes. Un excelente diagnóstico es la medida de la distancia desarrollada por Dennis R. Cook.

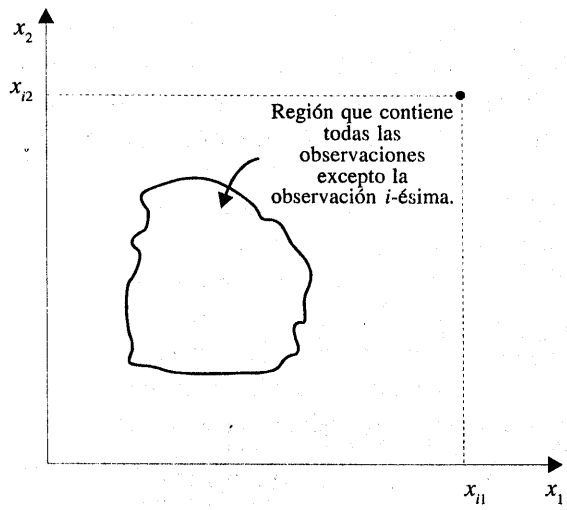


Figura 11-9 Un punto que está apartado en el espacio x.

Esta es una medida del cuadrado de la distancia entre la estimación de mínimos cuadrados usual de $\boldsymbol{\beta}$ basada en las n observaciones y la estimación obtenida cuando se elimina el punto i, digamos, $\hat{\boldsymbol{\beta}}_{(i)}$. La **medida de la distancia de Cook** es

$$D_i = \frac{\left(\hat{\boldsymbol{\beta}}_{(i)} - \hat{\boldsymbol{\beta}}\right)\mathbf{X}'\mathbf{X}\left(\hat{\boldsymbol{\beta}}_{(i)} - \hat{\boldsymbol{\beta}}\right)}{p\hat{\sigma}^2} \qquad i = 1, 2, \ldots, n$$

Evidentemente, si el punto i es influyente, su eliminación dará como resultado $\hat{\boldsymbol{\beta}}_{(i)}$ cambiando considerablemente el valor $\hat{\boldsymbol{\beta}}$. Por tanto, un valor grande de D_i implica que el punto i es influyente. El cálculo real del estadístico D_i se hace usando

$$D_i = \frac{r_i^2}{p}\frac{h_{ii}}{(1 - h_{ii})} \qquad i = 1, 2, \ldots, n \qquad (11\text{-}42)$$

En la ecuación 11-42 se observa que D_i consta del cuadrado del residual studentizado, lo cual refleja qué tan bien se ajusta el modelo a la observación i, y_i [recuérdese que $r_i = e_i / \sqrt{\hat{\sigma}^2 (1 - h_{ii})}$], y de un componente que mide qué tan apartado se encuentra ese punto del resto de los datos [$h_{ii}/(1 - h_{ii})$ es una medida de la distancia del punto i del centroide de los $n - 1$ puntos restantes]. Un valor $D_i > 1$ indicaría que el punto es influyente. Cualquiera de los componentes de D_i (o ambos) puede contribuir a un valor grande.

EJEMPLO 11-12

En la tabla 11-7 se enlistan los valores de las diagonales h_{ii} de la matriz gorro y la medida de la distancia de Cook D_i para los datos de la resistencia al desprendimiento del alambre adherido, del ejemplo 11-1. Para ilustrar los cálculos, considérese la primera observación:

$$
\begin{aligned}
D_1 &= \frac{r_1^2}{p} \cdot \frac{h_{11}}{(1-h_{11})} \\
&= \frac{\left[e_1 / \sqrt{MS_E\,(1-h_{11})}\right]^2}{p} \cdot \frac{h_{11}}{(1-h_{11})} \\
&= \frac{\left[1.57 / \sqrt{5.2352\,(1-0.1573)}\right]^2}{3} \cdot \frac{0.1573}{(1-0.1573)} \\
&= 0.035
\end{aligned}
$$

La medida de la distancia de Cook D_i no identifica ninguna observación potencialmente influyente en los datos, ya que ningún valor de D_i excede la unidad.

Tabla 11-7 Diagnóstico de influencia para los datos de la resistencia al desprendimiento del alambre adherido del ejemplo 11-12

Observaciones i	h_{ii}	Medida de la distancia de Cook D_i
1	0.1573	0.035
2	0.1116	0.012
3	0.1419	0.060
4	0.1019	0.021
5	0.0418	0.024
6	0.0749	0.007
7	0.1181	0.036
8	0.1561	0.020
9	0.1280	0.160
10	0.0413	0.001
11	0.0925	0.013
12	0.0526	0.001
13	0.0820	0.001
14	0.1129	0.003
15	0.0737	0.187
16	0.0879	0.001
17	0.2593	0.565
18	0.2929	0.155
19	0.0962	0.018
20	0.1473	0.000
21	0.1296	0.052
22	0.1358	0.028
23	0.1824	0.002
24	0.1091	0.040
25	0.0729	0.000

EJERCICIOS DE LA SECCIÓN 11-8

11-36. Considérese el modelo de regresión para los datos de la NFL del ejercicio 11-4.

a) ¿Qué proporción de la variabilidad total está explicada por este modelo?

b) Construya una gráfica de probabilidad normal de los residuales. ¿A qué conclusión puede llegarse a partir de esta gráfica?

c) Grafique los residuales contra $\hat{y}$ y contra cada una de las variables regresoras, y comente la adecuación del modelo.

d) ¿Existen puntos influyentes en estos datos?

11-37. Considérense los datos del rendimiento de gasolina del ejercicio 11-5.

a) ¿Qué proporción de la variabilidad total está explicada por este modelo?

b) Construya una gráfica de probabilidad normal de los residuales y comente el supuesto de normalidad.

c) Grafique los residuales contra $\hat{y}$ y contra cada una de las variables regresoras. Comente estas gráficas de los residuales.

d) Calcule la distancia de Cook para las observaciones de este conjunto de datos. ¿Alguna de las observaciones es influyente?

11-38. Considérense los datos del consumo de energía eléctrica del ejercicio 11-6.

a) Calcule R^2 para este modelo. Interprete esta cantidad.

b) Grafique los residuales contra $\hat{y}$. Interprete esta gráfica.

c) Construya una gráfica de probabilidad normal de los residuales y comente el supuesto de normalidad.

11-39. Considérense los datos del desgaste del rodamiento del ejercicio 11-7.

a) Encuentre el valor de R^2 cuando se usan las variables regresoras x_1 y x_2 en el modelo.

b) ¿Qué ocurre con el valor de R^2 cuando se agrega un término de interacción x_1x_2 al modelo? ¿Esto implica necesariamente que agregar el término de interacción es una buena idea?

11-40. Para el modelo de regresión de los datos de la resistencia al desprendimiento del alambre adherido del ejercicio 11-8.

a) Grafique los residuales contra $\hat{y}$ y contra las variables regresoras usadas en el modelo. ¿Qué información proporcionan estas gráficas?

b) Construya una gráfica de probabilidad normal de los residuales. ¿Hay alguna razón para dudar del supuesto de normalidad de este modelo?

c) ¿Hay algún indicio de una observación influyente en los datos?

11-41. Considérense los datos de la HFE del semiconductor del ejercicio 11-9.

a) Grafique los residuales de este modelo contra $\hat{y}$. Comente la información que proporciona esta gráfica.

b) ¿Cuál es el valor de R^2 para este modelo?

c) Ajuste de nuevo el modelo utilizando log HFE como variable de respuesta.

d) Grafique los residuales contra los valores predichos de log HFE para el modelo del inciso c) anterior. ¿Proporciona esto alguna información acerca de cuál de los modelos es preferible?

e) Grafique los residuales del modelo del inciso d) contra la variable regresora x_3. Comente esta gráfica.

f) Ajuste de nuevo el modelo para log HFE usando x_1, x_2 y $1/x_3$ como variables regresoras. Comente el efecto de este cambio en el modelo.

11-42. Considérese el modelo de regresión para los datos del tratamiento térmico del ejercicio 11-10.

a) Calcule el porcentaje de la variabilidad que explica este modelo.

b) Construya una gráfica de probabilidad normal para los residuales. Comente el supuesto de normalidad.

c) Grafique los residuales contra $\hat{y}$ e interprete la gráfica.

d) Calcule la distancia de Cook para cada observación y proponga una interpretación de este estadístico.

11-43. En el ejercicio 11-24 se ajustó un modelo a la variable de respuesta PITCH en los datos

del tratamiento térmico del ejercicio 11-10 usando las nuevas variables regresoras x_1 = SOAKTIME × SOAKPCT y x_2 = DIFFTIME × DIFFPCT.

a) Calcule el valor de R^2 para este modelo y compárelo con el valor de R^2 del modelo original del ejercicio 11-10. ¿Ofrece esto alguna información acerca de cuál de los modelos es preferible?

b) Grafique los residuales de este modelo contra $\hat{y}$ en una escala de probabilidad normal. Comente la adecuación del modelo.

c) Encuentre los valores de la medida de la distancia de Cook. ¿Hay alguna observación con una influencia excepcional?

11-44. Considérese el modelo de regresión para los datos de la NHL del ejercicio 11-11.

a) Ajuste un modelo usando "Puntos" como única variable regresora.

b) ¿Cuánta variabilidad está explicada por este modelo?

c) Grafique los residuales contra $\hat{y}$ y comente la adecuación del modelo.

d) Grafique los residuales contra "PPG", los tantos anotados en una jugada de poder. ¿Esta gráfica indica que el modelo se mejoraría si se incluyera esta variable?

11-45. Es común usar los elementos de la diagonal de la matriz gorro para denotar el **brazo de palanca**, es decir, un punto que es excepcional en su localización en el espacio x y que puede ser influyente. En general, al punto i se le llama **punto de brazo de palanca** si su diagonal gorro h_{ii} excede $2p/n$, que es el doble del tamaño promedio de todas las diagonales gorro. Recuérdese que $p = k + 1$.

a) La tabla 11-7 contiene la diagonal gorro para los datos de la resistencia al desprendimiento de un alambre adherido usados en el ejemplo 11-1. Encuetre el tamaño promedio de estos elementos.

b) Con base en el criterio anterior, ¿hay alguna observación que sea un punto de brazo de palanca en este conjunto de datos?

11-9 MODELOS DE REGRESIÓN POLINOMIAL

El modelo lineal $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ es un modelo general que puede usarse para ajustar cualquier relación que sea **lineal en los parámetros desconocidos** $\boldsymbol{\beta}$. Esto incluye la importante clase de los **modelos de regresión polinomiales**. Por ejemplo, el polinomio de segundo grado en una variable

$$Y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \varepsilon \tag{11-43}$$

y el polinomio de segundo grado en dos variables

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon \tag{11-44}$$

son modelos de regresión lineales.

Los modelos de regresión polinomiales se usan ampliamente cuando la respuesta es curvilínea, debido a que pueden aplicarse los principios generales de la regresión múltiple. En el ejemplo siguiente se ilustran algunos de los tipos de análisis que pueden realizarse.

EJEMPLO 11-13

Los paneles laterales del interior de un aeroplano se hacen en una prensa de 1 500 toneladas. El costo unitario de fabricación varía con el tamaño del lote de producción. Los datos que se presentan a continuación muestran el costo promedio por unidad (en cientos de dólares) de este

producto (y) y el tamaño del lote de producción (x). El diagrama de dispersión, ilustrado en la figura 11-10, indica que un polinomio de segundo orden puede ser apropiado.

y	1.81	1.70	1.65	1.55	1.48	1.40	1.30	1.26	1.24	1.21	1.20	1.18
x	20	25	30	35	40	50	60	65	70	75	80	90

Se ajustará el modelo

$$Y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \varepsilon$$

El vector $\mathbf{y}$, la matriz $\mathbf{X}$ y el vector $\boldsymbol{\beta}$ son los siguientes:

$$\mathbf{y} = \begin{bmatrix} 1.81 \\ 1.70 \\ 1.65 \\ 1.55 \\ 1.48 \\ 1.40 \\ 1.30 \\ 1.26 \\ 1.24 \\ 1.21 \\ 1.20 \\ 1.18 \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & 20 & 400 \\ 1 & 25 & 625 \\ 1 & 30 & 900 \\ 1 & 35 & 1\,225 \\ 1 & 40 & 1\,600 \\ 1 & 50 & 2\,500 \\ 1 & 60 & 3\,600 \\ 1 & 65 & 4\,225 \\ 1 & 70 & 4\,900 \\ 1 & 75 & 5\,625 \\ 1 & 80 & 6\,400 \\ 1 & 90 & 8\,100 \end{bmatrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_{11} \end{bmatrix}$$

Al resolver las ecuaciones normales $\mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{X'y}$ se obtiene el modelo ajustado

$$\hat{y} = 2.19826629 - 0.0252236x + 0.00012507x^2$$

La prueba de la significación de la regresión se muestra en la tabla 11-8. Puesto que $f_0 = 2\,171.07$ es significativo en el 1%, se concluye que al menos uno de los parámetros β_1 y β_{11} es diferente de cero. Además, las pruebas estándares de la adecuación del modelo no revelan ningún comportamiento extraño, y se concluiría que éste es un modelo razonable para los datos del costo de los paneles laterales.

Al ajustar polinomios, por lo general se preferirá usar el modelo de grado menor que sea consistente con los datos. En este ejemplo, sería razonable investigar la posibilidad de sacar el término cuadrático del modelo. Es decir, querría probarse

$$H_0 : \beta_{11} = 0$$
$$H_1 : \beta_{11} \neq 0$$

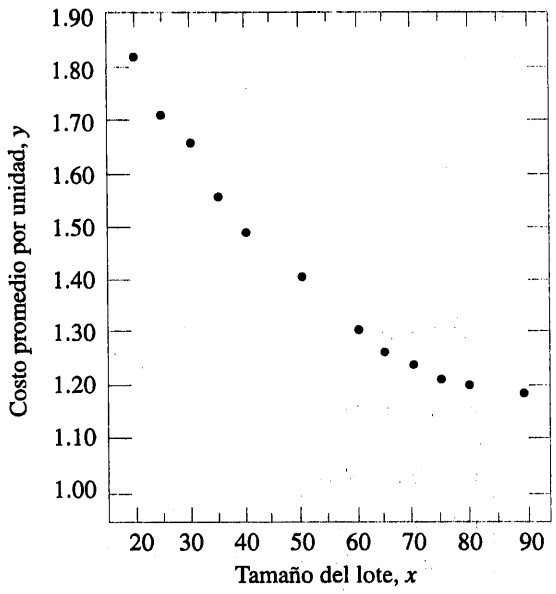


Figura 11-10 Datos del ejemplo 11-13.

Puede usarse la prueba general de significación de la regresión para probar esta hipótesis. Es necesario determinar la "suma extra de los cuadrados" debida a β_{11}, o

$$SS_R(\beta_{11}|\beta_1,\beta_0) = SS_R(\beta_1,\beta_{11}|\beta_0) - SS_R(\beta_1|\beta_0)$$

Por la tabla 11-9, la suma de los cuadrados $SS_R(\beta_{11},\beta_1|\beta_0) = 0.5254$. Para encontrar $SS_R(\beta_1|\beta_0)$, se ajusta un modelo de regresión lineal simple a los datos originales, de donde se obtiene

$$\hat{y} = 1.90036320 - 0.00910056x$$

Es fácil verificar que la suma de los cuadrados de la regresión para este modelo es

$$SS_R(\beta_1|\beta_0) = 0.4942$$

Tabla 11-8 Prueba de la significación de la regresión para el modelo de segundo orden del ejemplo 11-13

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Regresión	0.5254	2	0.262700	2 171.07	5.18E-15
Error	0.0011	9	0.000121		
Total	0.5265	11			

Tabla 11-9 Análisis de varianza del ejemplo 11-13, indicándose la prueba de H_0: $\beta_{11} = 0$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Regresión	$SS_R(\beta_1, \beta_{11} \mid \beta_0) = 0.5254$	2	0.262700	2 171.07	5.18E-15
Lineal	$SS_R(\beta_1 \mid \beta_0) = 0.4942$	1	0.494200	4 084.30	1.17E-15
Cuadrática	$SS_R(\beta_{11} \mid \beta_0, \beta_1) = 0.0312$	1	0.031200	258.18	5.51E-9
Error	0.0011	9	0.00121		
Total	0.5265	11			

Por lo tanto, la suma extra de los cuadrados debida a β_{11}, dado que β_1 y β_0 están en el modelo, es

$$\begin{aligned} SS_R(\beta_{11} \mid \beta_1, \beta_0) &= SS_R(\beta_1, \beta_{11} \mid \beta_0) - SS_R(\beta_1 \mid \beta_0) \\ &= 0.5254 - 0.4942 \\ &= 0.312 \end{aligned}$$

El análisis de varianza, con la prueba de H_0: $\beta_{11} = 0$ incorporada en el proceso, se muestra en la tabla 11-9. Obsérvese que el término cuadrático contribuye de manera significativa al modelo.

11-10 VARIABLES INDICADORAS

Los modelos de regresión presentados en las secciones anteriores se basan en variables **cuantitativas**, es decir, variables que se miden en una escala numérica. Por ejemplo, variables tales como la temperatura, la presión, la distancia y el voltaje son variables cuantitativas. Ocasionalmente, es necesario incorporar variables **cualitativas** en un modelo de regresión. Por ejemplo, suponga que una de las variables de un modelo de regresión es el operador que está asociado con cada observación y_i. Suponga que sólo participan dos operadores. Quizá quieran asignarse niveles diferentes a los dos operadores para tomar en consideración la posibilidad de que cada operador puede tener un efecto diferente sobre la variable de respuesta.

El método común para tomar en consideración los diferentes niveles de una variable cualitativa es usando **variables indicadoras**. Por ejemplo, para introducir el efecto de dos operadores diferentes en un modelo de regresión, podría definirse una variable indicadora como sigue:

$$x = \begin{cases} 0 & \text{si la observación es del operador 1} \\ 1 & \text{si la observación es del operador 2} \end{cases}$$

En general, una variable cualitativa con t niveles puede modelarse con $t - 1$ variables indicadoras, a las que se asigna el valor cero o uno. Por tanto, si hay *tres* operadores, los diferentes niveles se tomarán en consideración con *dos* variables indicadoras definidas como sigue:

x_1	x_2	
0	0	si la observación es del operador 1
1	0	si la observación es del operador 2
0	1	si la observación es del operador 3

Las variables indicadoras también se conocen como variables **mudas**. En el ejemplo siguiente (tomado de Montgomery y Peck) se ilustran algunos de los usos de las variables indicadoras; para otras aplicaciones, véase Montgomery y Peck.

EJEMPLO 11-14

Un ingeniero mecánico investiga el acabado superficial de las piezas metálicas fabricadas en un torno y su relación con la velocidad (en RPM) del torno. Los datos se presentan en la tabla 11-10. Obsérvese que los datos se han recabado usando dos tipos diferentes de herramientas de corte. Puesto que es probable que el tipo de herramienta de corte afecte el acabado superficial, se ajustará el modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Tabla 11-10 Datos del acabado superficial del ejemplo 11-14

Número de observación, i	Acabado superficial y_i	RPM	Tipo de herramienta de corte
1	45.44	225	302
2	42.03	200	302
3	50.10	250	302
4	48.75	245	302
5	47.92	235	302
6	47.79	237	302
7	52.26	265	302
8	50.52	259	302
9	45.58	221	302
10	44.78	218	302
11	33.50	224	416
12	31.23	212	416
13	37.52	248	416
14	37.13	260	416
15	34.70	243	416
16	33.92	238	416
17	32.13	224	416
18	35.47	251	416
19	33.49	232	416
20	32.29	216	416

donde Y es el acabado superficial, x_1 es la velocidad del torno en RPM, y x_2 es una variable indicadora que denota el tipo de herramienta de corte usada; es decir,

$$x_2 = \begin{cases} 0, & \text{para la herramienta tipo 302} \\ 1, & \text{para la herramienta tipo 416} \end{cases}$$

Los parámetros de este modelo pueden interpretarse con facilidad. Si $x_2 = 0$, entonces el modelo es

$$Y = \beta_0 + \beta_1 x_1 + \varepsilon$$

que es un modelo de línea recta con pendiente β_1 y ordenada al origen β_0. Sin embargo, si $x_2 = 1$, entonces el modelo es

$$Y = \beta_0 + \beta_1 x_1 + \beta_2(1) + \varepsilon = \beta_0 + \beta_2 + \beta_1 x_1 + \varepsilon$$

que es un modelo de linea recta con pendiente β_1 y ordenada al origen $\beta_0 + \beta_2$. Por tanto, el modelo $Y = \beta_0 + \beta_1 x + \beta_2 x_2 + \varepsilon$ implica que el acabado superficial se relaciona linealmente con la velocidad del torno y que la pendiente β_1 no depende del tipo de herramienta de corte usada. Sin embargo, el tipo de herramienta de corte sí afecta a la ordenada al origen, y β_2 indica el cambio en la ordenada al origen que se asocia con un cambio en el tipo de herramienta de 302 a 416.

La matriz **X** y el vector **y** para este problema son los siguientes:

$$\mathbf{X} = \begin{bmatrix} 1 & 225 & 0 \\ 1 & 200 & 0 \\ 1 & 250 & 0 \\ 1 & 245 & 0 \\ 1 & 235 & 0 \\ 1 & 237 & 0 \\ 1 & 265 & 0 \\ 1 & 259 & 0 \\ 1 & 221 & 0 \\ 1 & 218 & 0 \\ 1 & 224 & 1 \\ 1 & 212 & 1 \\ 1 & 248 & 1 \\ 1 & 260 & 1 \\ 1 & 243 & 1 \\ 1 & 238 & 1 \\ 1 & 224 & 1 \\ 1 & 251 & 1 \\ 1 & 232 & 1 \\ 1 & 216 & 1 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 45.44 \\ 42.03 \\ 50.10 \\ 48.75 \\ 47.92 \\ 47.79 \\ 52.26 \\ 50.52 \\ 45.58 \\ 44.78 \\ 33.50 \\ 31.23 \\ 37.52 \\ 37.13 \\ 34.70 \\ 33.92 \\ 32.13 \\ 35.47 \\ 33.49 \\ 32.29 \end{bmatrix}$$

El modelo ajustado es

$$\hat{y} = 14.27620 + 0.14115x_1 - 13.28020x_2$$

En la tabla 11-11 se muestra el análisis de varianza para este modelo. Obsérvese que la hipótesis H_0: $\beta_1 = \beta_2 = 0$ (la significación de la regresión) se rechazaría en cualquier nivel razonable de significación debido a que el valor P es muy pequeño. Esta tabla también contiene las sumas de los cuadrados

$$\begin{aligned} SS_R &= SS_R(\beta_1\beta_2|\beta_0) \\ &= SS_R(\beta_1|\beta_0) + SS_R(\beta_2|\beta_1,\beta_0) \end{aligned}$$

por lo que puede hacerse la prueba de la hipótesis H_0: $\beta_2 = 0$. Puesto que esta hipótesis también es rechazada, se concluye que el tipo de herramienta tiene un efecto sobre el acabado superficial.

También es posible usar variables indicadoras para investigar si el tipo de herramienta afecta tanto a la pendiente como a la ordenada al origen. Sea el modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

donde x_2 es la variable indicadora. Entonces, si se usa la herramienta tipo 302, $x_2 = 0$, y el modelo es

$$Y = \beta_0 + \beta_1 x_1 + \varepsilon$$

Si se usa la herramienta tipo 416, $x_2 = 1$, y el modelo es

$$\begin{aligned} Y &= \beta_0 + \beta_1 x_1 + \beta_2 + \beta_3 x_1 + \varepsilon \\ &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3)x_1 + \varepsilon \end{aligned}$$

Obsérvese que β_2 es el cambio en la ordenada al origen y que β_3 es el cambio en la pendiente producido por un cambio en el tipo de herramienta.

Tabla 11-11 Análisis de varianza del ejemplo 11-14

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Regresión	1 012.0595	2	506.0297	1 103.69	1.02 E-18
$SS_R(\beta_1 \mid \beta_0)$	130.6091	1	130.6091	284.87	4.70 E-12
$SS_R(\beta_2 \mid \beta_1, \beta_0)$	881.4504	1	881.4504	1 922.52	6.24 E-19
Error	7.7943	17	0.4508		
Total	1 019.8538	19			

Otro método para analizar estos datos es ajustar modelos de regresión separados a los datos para cada tipo de herramienta. Sin embargo, el enfoque de la variable indicadora tiene varias ventajas. Primera, únicamente debe ajustarse un modelo de regresión. Segunda, al agrupar los datos de ambos tipos de herramienta, se obtienen más grados de libertad para el error. Tercera, las pruebas de ambas hipótesis para los parámetros β_2 y β_3 son tan sólo casos especiales de la prueba general de significación de la regresión.

EJERCICIOS DE LAS SECCIONES 11-9 Y 11-10

11-46. En un artículo titulado "Método para mejorar la precisión del análisis de regresión polinomial" de *Journal of Quality Technology* se reportaron los siguientes datos para y = resistencia al corte final de un compuesto de hule (psi) y x = temperatura del curado (°F).

y	770	800	840	810
x	280	284	292	295
y	735	640	590	560
x	298	305	308	315

a) Ajuste un polinomio de segundo grado a estos datos.

b) Pruebe la significación de la regresión usando $\alpha = 0.05$.

c) Pruebe la hipótesis de que $\beta_{11} = 0$ usando $\alpha = 0.05$.

d) Calcule los residuales del inciso a) y úselos para evaluar la adecuación del modelo.

11-47. Considérense los siguientes datos, que resultan de un experimento para determinar el efecto de x = tiempo de prueba en horas a una temperatura particular sobre y = cambio en la viscosidad del aceite.

y	−1.42	−1.39	−1.55	−1.89	−2.43
x	.25	.50	.75	1.00	1.25
y	−3.15	−4.05	−5.15	−6.43	−7.89
x	1.50	1.75	2.00	2.25	2.50

a) Ajuste un polinomio de segundo orden a los datos.

b) Pruebe la significación de la regresión usando $\alpha = 0.05$.

c) Pruebe la hipótesis de que $\beta_{11} = 0$ usando $\alpha = 0.05$.

d) Calcule los residuales del inciso a) y úselos para evaluar la adecuación del modelo.

11-48. Cuando se ajustan modelos de regresión polinomiales, con frecuencia se resta $\bar{x}$ de cada valor x para producir una variable regresora "centrada" $x' = x - \bar{x}$. Con esto se reducen los efectos de dependencia entre los términos del modelo y con frecuencia se llega a estimaciones más precisas de los coeficientes de regresión. Usando los datos del ejercicio 11-46, ajuste el modelo $Y = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2 + \varepsilon$. Use los resultados para estimar los coeficientes del modelo no centrado $Y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \varepsilon$.

11-49. Suponga que se usa la variable estandarizada $x' = (x - \bar{x})/s_x$, donde s_x es la desviación estándar de x, para construir un modelo de regresión polinomial. Usando los datos del ejercicio 11-46 y el enfoque de las variables estandarizadas, ajuste el modelo $Y = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2 + \varepsilon$.

a) ¿Qué valor de y se predice cuando x = 285°F?

b) Estime los coeficientes de regresión del modelo no estandarizado $Y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \varepsilon$.

c) ¿Qué puede decirse acerca de la relación entre SS_E y R^2 para los modelos estandarizado y no estandarizado?

d) Suponga que se usa $y' = (y - \bar{y})/s_y$ en el modelo junto con x'. Ajustar el modelo y comente la relación entre SS_E y R^2 para el modelo estandarizado y el no estandarizado.

11-50. Los datos que se presentan abajo se recabaron durante un experimento para determinar el cambio en la eficiencia del impul-

so (y, en por ciento) cuando el ángulo de divergencia de la nariz de un cohete (x) cambia.

y	24.60	24.71	23.90	39.50	39.60	57.12
x	4.0	4.0	4.0	5.0	5.0	6.0
y	67.11	67.24	67.15	77.87	80.11	84.67
x	6.5	6.5	6.75	7.0	7.1	7.3

a) Ajuste un modelo de segundo orden a estos datos.
b) Pruebe la significación de la regresión y la falta de ajuste usando $\alpha = 0.05$.
c) Pruebe la hipótesis de que $\beta_{11} = 0$ usando $\alpha = 0.05$.
d) Grafique los residuales y comente la adecuación del modelo.
e) Ajuste un modelo cúbico y pruebe la significación del término cúbico usando $\alpha = 0.05$.

11-51. En un artículo de *Journal of Pharmaceuticals Sciences* (vol. 80, pp. 971-977) se presenta datos de la solubilidad observada de la fracción molar de un soluto a una temperatura constante y los parámetros de Hansen de la solubilidad parcial, de dispersión, bipolaridad y de enlace de hidrógeno. Los datos se presentan en la siguiente tabla, donde y es el logaritmo negativo de la solubilidad de la fracción molar, x_1 es la solubilidad parcial de dispersión, x_2 es la solubilidad parcial bipolar, y x_3 es la solubilidad parcial de enlace de hidrógeno.

a) Ajuste el modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \varepsilon$.
b) Pruebe la significación de la regresión usando $\alpha = 0.05$.
c) Grafique los residuales y comente la adecuación del modelo.
d) Use el método de la suma extra de los cuadrados para probar la contribución de los términos de segundo orden usando $\alpha = 0.05$.

Número de observación	y	x_1	x_2	x_3
1	0.22200	7.3	0.0	0.0
2	0.39500	8.7	0.0	0.3
3	0.42200	8.8	0.7	1.0
4	0.43700	8.1	4.0	0.2
5	0.42800	9.0	0.5	1.0
6	0.46700	8.7	1.5	2.8
7	0.44400	9.3	2.1	1.0
8	0.37800	7.6	5.1	3.4
9	0.49400	10.0	0.0	0.3
10	0.45600	8.4	3.7	4.1
11	0.45200	9.3	3.6	2.0
12	0.11200	7.7	2.8	7.1
13	0.43200	9.8	4.2	2.0
14	0.10100	7.3	2.5	6.8
15	0.23200	8.5	2.0	6.6
16	0.30600	9.5	2.5	5.0
17	0.09230	7.4	2.8	7.8
18	0.11600	7.8	2.8	7.7
19	0.07640	7.7	3.0	8.0
20	0.43900	10.3	1.7	4.2
21	0.09440	7.8	3.3	8.5
22	0.11700	7.1	3.9	6.6
23	0.07260	7.7	4.3	9.5
24	0.04120	7.4	6.0	10.9
25	0.25100	7.3	2.0	5.2
26	0.00002	7.6	7.8	20.7

11-52. Considérense los datos del rendimiento de gasolina del ejercicio 11-5.

a) Comente cómo se modelaría la información acerca del tipo de transmisión del automóvil.
b) Ajuste un modelo de regresión a las millas por galón, usando el cilindraje del motor, los caballos de fuerza, y el tipo de transmisión del automóvil como variables regresoras.
c) ¿Hay evidencia de que el tipo de transmisión afecta el rendimiento de la gasolina?

11-53. Considérense los datos del tipo de herramienta del ejemplo 11-14. Pruebe la hipótesis de que se necesitan dos modelos de regresión diferentes (con pendientes y ordenadas al origen diferentes) para modelar adecuadamente los datos. Use variables indicadoras para resolver este ejercicio.

11-11 SELECCIÓN DE VARIABLES EN REGRESIONES MÚLTIPLES

11-11.1 Problema de la construcción del modelo

Un problema importante en muchas aplicaciones del análisis de regresión incluye la selección del conjunto de variables regresoras o independientes que van a usarse en el modelo. En ocasiones experiencias previas o consideraciones de la teoría fundamental pueden ayudar al analista a especificar el conjunto de variables regresoras que deberán usarse en una situación particular. Sin embargo, por lo general el problema consiste en seleccionar un conjunto apropiado de variables regresoras de un conjunto que muy probablemente incluye todas las variables importantes, aunque se tiene la seguridad de que no todas estas variables regresoras candidatas son necesarias para modelar adecuadamente la variable de respuesta Y.

En una situación como ésta, el interés se encuentra en depurar las variables candidatas para obtener un modelo de regresión que contenga el "mejor" subconjunto de variables regresoras. Lo ideal sería que el modelo final incluyera suficientes variables regresoras para que en el uso proyectado del modelo (predicción, por ejemplo) tuviera un desempeño satisfactorio. Por otra parte, para mantener los costos de mantenimiento del modelo en un mínimo y para hacer que el modelo sea fácil de usar, sería conveniente que el modelo usara el menor número posible de variables regresoras. Al término medio entre estos objetivos opuestos suele llamársele encontrar la "mejor" ecuación de regresión. Sin embargo, en la mayoría de los problemas, ningún modelo de regresión particular es "mejor" en términos de los diferentes criterios de evaluación que se han propuesto. Por lo general, se necesitará mucho criterio y experiencia con el sistema que se está modelando para seleccionar un conjunto apropiado de variables regresoras para una ecuación de regresión.

Ningún algoritmo particular producirá siempre una buena solución al problema de la selección de las variables. La mayoría de los procedimientos disponibles actualmente son técnicas de búsqueda, y para que su desempeño sea satisfactorio requieren la interacción y el criterio del analista. Se revisan ahora brevemente algunas de las técnicas de selección de variables de uso más común.

11-11.2 Procedimientos computacionales para la selección de variables

Se supone que hay K variables regresoras candidatas $x_1, x_2, ..., x_K$ y una sola variable de respuesta y. Todos los modelos incluirán un término β_0 para la ordenada al origen, por lo que el modelo con *todas* las variables incluidas tendrá $K + 1$ términos. Además, la forma funcional de cada variable candidata (por ejemplo, $x_1 = 1/x$, $x_2 = \ln x$, etc.) es correcta.

Todas las regresiones posibles

Este enfoque requiere que el analista ajuste todas las ecuaciones de regresión que incluyen una variable candidata, todas las ecuaciones de regresión que incluyen dos variables candidatas, y así sucesivamente. Después, estas ecuaciones se evalúan de acuerdo con algún criterio adecuado para seleccionar el "mejor" modelo de regresión. Si hay K variables regresoras candidatas, será necesario examinar un total de 2^K ecuaciones. Por ejemplo, si $K = 4$, hay $2^4 = 16$ ecuaciones de regresión posibles; mientras que si $K = 10$, hay $2^{10} = 1\,024$ ecuaciones de regresión posibles.

Por tanto, el número de ecuaciones por examinar aumenta rápidamente conforme se incrementa el número de variables candidatas.

Pueden usarse varios criterios para evaluar y comparar los diferentes modelos de regresión obtenidos. Un criterio de uso común se basa en el coeficiente de determinación múltiple. Sea que R_p^2 denote el coeficiente de determinación para un modelo de regresión con p términos, es decir, con $p - 1$ variables regresoras y un término para la ordenada al origen (obsérvese que $p \leq K + 1$). En cuanto a los cálculos por realizar, se tiene

$$R_p^2 = \frac{SS_R(p)}{SS_T} = 1 - \frac{SS_E(p)}{SS_T} \tag{11-45}$$

donde $SS_R(p)$ y $SS_E(p)$ denotan la suma de cuadrados de la regresión y la suma de cuadrados del error, respectivamente, para un modelo con p variables. Entonces, R_p^2 se incrementa cuando p se incrementa y es un máximo cuando $p = K + 1$. Por lo tanto, el analista agrega variables al modelo hasta el punto en que una variable adicional ya no es útil porque resulta en un incremento reducido de R_p^2. El enfoque general se ilustra en la figura 11-11, donde se presenta una gráfica hipotética de R_p^2 contra p. De manera típica, el analista examina una gráfica como ésta y elige el número de variables regresoras del modelo como el punto donde es evidente el "codo" de la curva. Desde luego, hacer esta elección requiere criterio por parte del analista.

Un segundo criterio es considerar el cuadrado medio de error de una ecuación con p variables, denotado como $MS_E(p) = SS_E(p)/(n - p)$. En general, $MS_E(p)$ decrece cuando p se incrementa, aunque no necesariamente ocurre esto. Si la incorporación de una variable regresora al modelo con $p - 1$ términos no reduce la suma de cuadrados del error en una cantidad igual al cuadrado medio de error del modelo con los $p - 1$ términos originales, $MS_E(p)$ *se incrementará* debido a la pérdida de un grado de libertad para el error. Por lo tanto, un criterio lógico es elegir las variables regresoras de tal modo que $MS_E(p)$ sea un mínimo. De manera alternativa, puesto que $MS_E(p)$ por lo general es relativamente plano en la vecindad del mínimo, podría escogerse un modelo de tal modo que al agregarse más variables regresoras al modelo se produzcan tan sólo reducciones pequeñas en $MS_E(p)$. El procedimiento general se ilustra en la figura 11-12.

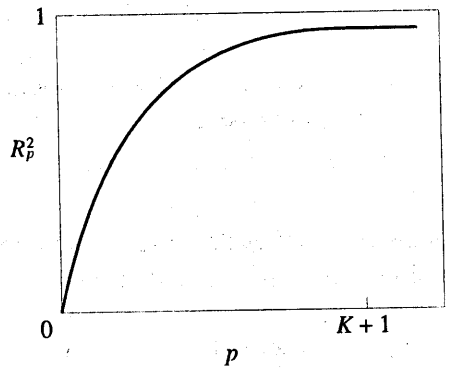


Figura 11-11 Gráfica de R_p^2 contra p.

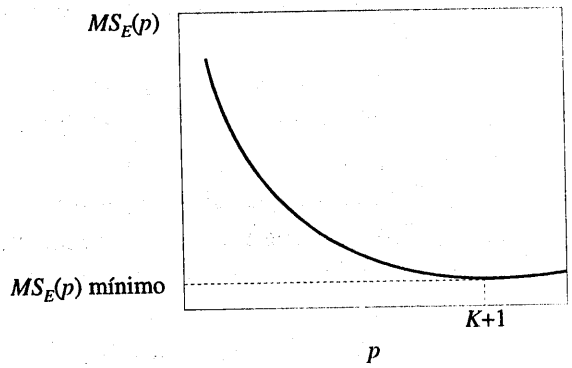


Figura 11-12 Gráfica de $MS_E(p)$ contra p.

Un tercer criterio es el estadístico C_p, que es una medida del cuadrado medio de error total del modelo de regresión. El cuadrado medio de error total estandarizado del modelo de regresión se define como

$$\begin{aligned}\Gamma_p &= \frac{1}{\sigma^2}\sum_{i=1}^{n} E\left[\hat{Y}_i - E(Y_i)\right]^2 \\ &= \frac{1}{\sigma^2}\left\{\sum_{i=1}^{n}\left[E(Y_i) - E(\hat{Y}_i)\right]^2 + \sum_{i=1}^{n} V(\hat{Y}_i)\right\} \\ &= \frac{1}{\sigma^2}\left[(\text{sesgo})^2 + \text{varianza}\right]\end{aligned}$$

Se usa el cuadrado medio de error del modelo *pleno* con los $K + 1$ términos como estimación de σ^2; es decir, $\hat{\sigma}^2 = MS_E(K + 1)$. Entonces un estimador de Γ_p es [véase Montgomery y Peck o Myers para los detalles]:

$$C_p = \frac{SS_E(p)}{\hat{\sigma}^2} - n + 2p \tag{11-46}$$

Si el modelo con p términos tiene un sesgo que se puede pasar por alto, entonces puede demostrarse que

$$E(C_p \mid \text{sesgo cero}) = p$$

Por lo tanto, los valores de C_p para cada modelo de regresión bajo consideración deberán evaluarse con respecto a p. Las ecuaciones de regresión que tienen sesgo que se puede omitir tendrán valores de C_p que están cerca de p, mientras las que tienen un sesgo significativo tendrán valores de C_p que son significativamente mayores que p. Entonces se escoge como la "mejor" ecuación de regresión un modelo con C_p *mínimo* o bien un modelo con C_p ligeramente mayor que no contenga un sesgo importante (es decir, $C_p \cong p$).

Otro criterio se basa en una modificación de R_p^2 que sirve para considerar el número de variables del modelo. A este estadístico se le llama R_p^2 **ajustado**, el cual se define como

$$\bar{R}_p^2 = 1 - \frac{n-1}{n-p}\left(1 - R_p^2\right) \tag{11-47}$$

Obsérvese que $\bar{R}_p^2$ puede decrecer cuando p se incrementa si el decremento en $(n-1)(1-\bar{R}_p^2)$ no es compensado por la pérdida de un grado de libertad en $n - p$. El experimentador por lo general

seleccionaría el modelo de regresión que tenga el valor máximo de $\bar{R}_p^2$. Sin embargo, obsérvese que esto es equivalente al modelo que minimiza $MS_E(p)$, ya que

$$\begin{aligned}\bar{R}_p^2 &= 1-\left(\frac{n-1}{n-p}\right)\left(1-R_p^2\right)\\ &= 1-\left(\frac{n-1}{n-p}\right)\frac{SS_E(p)}{SS_T}\\ &= 1-\left(\frac{n-1}{SS_T}\right)MS_E(p)\end{aligned}$$

EJEMPLO 11-15

Los datos de la tabla 11-12 son una ampliación del conjunto de datos del estudio de la resistencia al desprendimiento del alambre adherido del ejemplo 11-1. Ahora hay cuatro variables candidatas: la longitud del alambre (x_1), la altura de la matriz (x_2), la altura del poste (x_3), la altura del amarre (x_4).

En la tabla 11-13 se presentan los resultados de aplicar todas las regresiones posibles (excepto el modelo trivial con una sola ordenada al origen) a estos datos. En esta tabla también se

Tabla 11-12 Datos de la resistencia al desprendimiento del alambre adherido para el ejemplo 11-15

Número de observación	Resistencia al desprendimiento y	Longitud del alambre x_1	Altura de la matriz x_2	Altura del poste x_3	Altura del amarre x_4
1	9.95	2	50	1	1
2	24.45	8	110	1	1
3	31.75	11	120	2	1
4	35.00	10	550	2	2
5	25.02	8	295	1	1
6	16.86	4	200	1	1
7	14.38	2	375	1	1
8	9.60	2	52	1	1
9	24.35	9	100	1	1
10	27.50	8	300	2	1
11	17.08	4	412	2	2
12	37.00	11	400	3	2
13	41.95	12	500	3	3
14	11.66	2	360	1	1
15	21.65	4	205	2	2
16	17.89	4	400	2	1
17	69.00	20	600	4	4
18	10.30	1	585	1	1
19	34.93	10	540	2	1
20	46.59	15	250	3	2
21	44.88	15	290	3	1
22	54.12	16	510	3	3
23	56.63	17	590	2	2
24	22.13	6	100	2	1
25	21.15	5	400	1	1

Tabla 11-13 Todas las regresiones posibles para los datos del ejemplo 11-15

Número de variables del modelo	p	Variables del modelo	R_p^2	$SS_R(p)$	$SS_E(p)$	$MS_E(p)$	$\bar{R}_p^2$	C_p
1	2	x_2	.24291747	1 483.2406	4 622.7041	200.9871	.2100	1 479.93
1	2	x_4	.56007492	3 419.7865	2 686.1582	116.7895	.5409	851.16
1	2	x_3	.39830969	4 263.8404	1 842.1043	80.0915	.6852	577.11
1	2	x_1	.96395437	5 885.8521	220.0926	9.5692	.9624	50.46
2	3	x_2, x_4	.57158123	3 490.0434	2 615.9013	118.9046	.5326	830.35
2	3	x_2, x_3	.72162460	4 406.1999	1 699.7448	77.2611	.6963	532.88
2	3	x_3, x_4	.72232683	4 410.4877	1 695.4570	77.0662	.6971	531.49
2	3	x_1, x_3	.97220237	5 936.2139	169.7308	7.7150	.9697	36.11
2	3	x_1, x_2	.98113748	5 990.7712	115.1735	5.2352	.9794	18.40
2	3	x_1, x_4	.98302963	6 002.3246	103.6201	4.7100	.9815	14.64
3	4	x_2, x_3, x_4	.73299076	4 475.6010	1 630.3437	77.6354	.6948	512.35
3	4	x_1, x_3, x_4	.98327895	6 003.8469	102.0978	4.8618	.9809	16.15
3	4	x_1, x_2, x_3	.98517507	6 015.4245	90.5203	4.3105	.9831	12.39
3	4	x_1, x_2, x_4	.98961493	6 042.5340	63.4107	3.0196	.9881	3.59
4	5	x_1, x_2, x_3, x_4	.98991196	6 044.3477	61.5970	3.0799	.9879	5

muestran los valores de R_p^2, $\bar{R}_p^2$, $SS_R(p)$, $SS_E(p)$, $MS_E(p)$ y C_p, y en la figura 11-13 se presenta una gráfica de R_p^2 contra p. En términos del mejoramiento de R^2, se gana poco al pasar de un modelo de dos variables a uno de tres variables. En la figura 11-14 se muestra una gráfica del $MS_E(p)$ mínimo para cada subconjunto de tamaño p. Varios modelos tienen valores adecuados de $MS_E(p)$. El mejor modelo de dos variables es (x_1, x_2) o bien (x_1, x_4), y el mejor modelo de tres variables es (x_1, x_2, x_4). El valor mínimo de $MS_E(p)$ ocurre para el modelo de tres variables (x_1, x_2, x_4). Aun cuando otros modelos tienen valores relativamente pequeños de $MS_E(p)$, como (x_1, x_2, x_3) y (x_1, x_2), el modelo con las variables regresoras (x_1, x_2, x_4) es definitivamente superior con respecto al criterio del $MS_E(p)$. Obsérvese que, como sería de esperarse, este modelo también minimiza el R_p^2 ajustado. En la figura 11-15 se muestra una gráfica del estadístico C_p. El único modelo con $C_p \leq p$ es la ecuación de tres variables con las variables regresoras (x_1, x_2, x_4). Para ilustrar los cálculos, en esta ecuación se encontraría que

$$C_p = \frac{SS_E(p)}{\hat{\sigma}^2} - n + 2p$$
$$= \frac{63.4107}{3.0799} - 25 + 2(4) = 3.59$$

observándose que $\hat{\sigma}^2 = 3.0799$ se obtiene de la ecuación *plena* o *saturada* (x_1, x_2, x_3, x_4). Puesto que los demás modelos contienen un sesgo sustancial $(C_p > p)$, con base en el criterio C_p se concluiría que el mejor subconjunto de variables regresoras es (x_1, x_2, x_4). Puesto que este modelo también resulta en un $MS_E(p)$ mínimo y en un R_p^2 alto, se le seleccionaría como la "mejor" ecuación de regresión. El modelo final es

$$\hat{y} = 1.36707 + 2.53492x_1 + 0.0085x_2 + 2.59928x_4$$

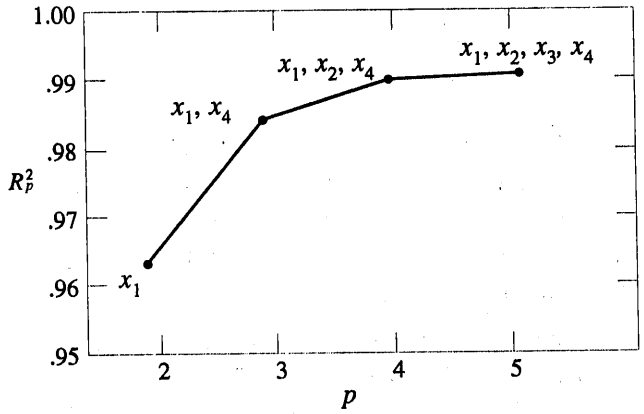


Figura 11-13 Gráfica de R_p^2 del ejemplo 11-15.

El enfoque de todas las regresiones posibles requiere un esfuerzo de cálculo sustancial, incluso cuando K es moderadamente pequeña. Sin embargo, si el analista está dispuesto a examinar algo menos que los coeficientes del modelo ajustado y todos los estadísticos del modelo asociados, es posible crear algoritmos para todas las regresiones posibles que producen menos información acerca de cada modelo pero que son más eficientes en lo que al cálculo se refiere. Por ejemplo, suponga que pudiera calcularse eficientemente sólo MS_E de cada modelo. Puesto que los modelos con MS_E grandes tienen pocas posibilidades de ser seleccionados como las mejores ecuaciones de regresión, entonces sólo tendrían que examinarse en detalle los modelos con valores pequeños de MS_E. Pueden usarse varios enfoques para desarrollar un algoritmo computacional eficiente desde el punto de vista de los cálculos, para todas las regresiones posibles. Para una revisión más amplia de estos algoritmos e ilustraciones de su uso, véase Montgomery y Peck y Myers. De manera típica, los problemas que tienen hasta 30 variables candidatas pueden resolverse fácilmente con este enfoque.

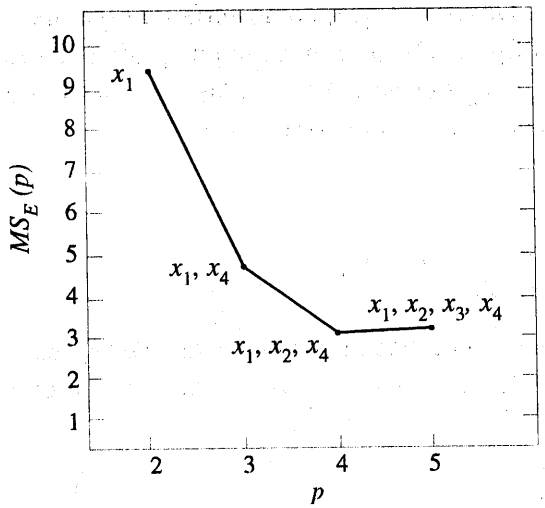


Figura 11-14 Gráfica de $MS_E(p)$ del ejemplo 11-15.

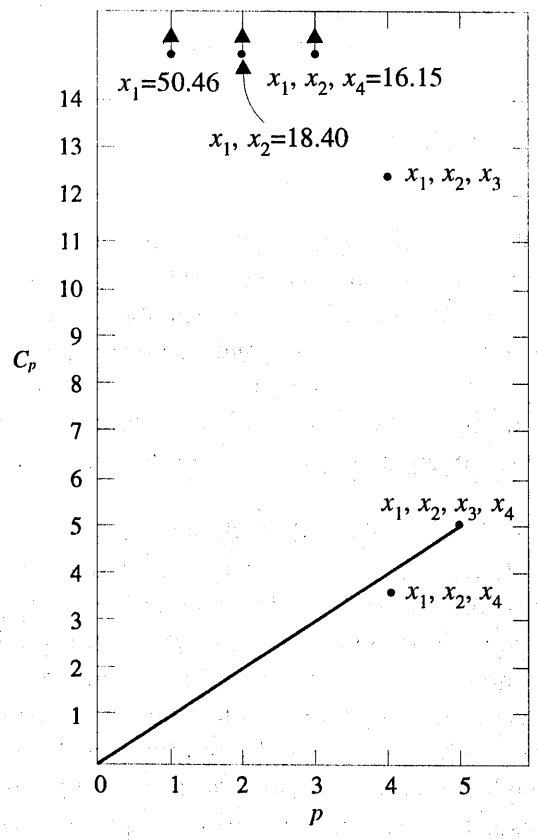


Figura 11-15 Gráfica de C_p del ejemplo 11-15.

Regresión por pasos

Probablemente la regresión por pasos sea la técnica de selección de variables de uso más generalizado. El procedimiento construye iterativamente una secuencia de modelos de regresión agregando o eliminando variables en cada paso. El criterio para agregar o eliminar una variable en cualquier paso suele expresarse en términos de una prueba F parcial. Sea f_{in} el valor de la variable aleatoria F, por agregar una variable al modelo, y sea f_{out} el valor de la variable aleatoria F por eliminar una variable del modelo. Debe tenerse $f_{in} \geq f_{out}$ y por lo general $f_{in} = f_{out}$.

La regresión por pasos se inicia formando un modelo de una variable utilizando la variable regresora que tiene la correlación más alta con la variable de respuesta Y. Se trata de la variable regresora que también producirá el estadístico F mayor. Por ejemplo, suponga que en este paso se selecciona x_1. En el segundo paso, se examinan las $K - 1$ variables candidatas restantes, y la variable para la que el estadístico F parcial

$$F_j = \frac{SS_R(\beta_j | \beta_1, \beta_0)}{MS_E(x_j, x_1)} \tag{11-48}$$

es un máximo se agrega a la ecuación, siempre que $f_j > f_{in}$. En la ecuación 11-48, $MS_E(x_j, x_1)$ denota el cuadrado medio de error del modelo que contiene tanto a x_1 como a x_j. Suponga que

este procedimiento indica que x_2 deberá agregarse al modelo. Entonces el algoritmo de regresión por pasos determina si la variable x_1 que se agregó en el primer paso deberá eliminarse. Esto se hace calculando el estadístico F

$$F_1 = \frac{SS_R(\beta_1|\beta_2, \beta_0)}{MS_E(x_1, x_2)} \tag{11-49}$$

Si el valor calculado $f_1 < f_{out}$, la variable x_1 se elimina; en caso contrario se conserva, y se intentaría agregar una variable regresora al modelo que contenga tanto a x_1 como a x_2.

En general, en cada paso se examina el conjunto de las variables regresoras restantes, y se agrega la variable regresora con el estadístico F parcial más grande, siempre que el valor observado de f exceda f_{in}. Después se calcula el estadístico F parcial para cada variable regresora del modelo, y la variable regresora con el valor menor de F observado se elimina si la f observada < f_{out}. El procedimiento continúa hasta que no pueden agregarse o eliminarse otras variables regresoras del modelo.

La regresión por pasos casi siempre se lleva a cabo utilizando un programa de computadora. El analista controla el procedimiento eligiendo f_{in} y f_{out}. Algunos programas de computadora de regresiones por pasos requieren que se especifiquen los valores numéricos de f_{in} y f_{out}. Puesto que el número de grados de libertad de MS_E depende del número de variables del modelo, el cual cambia de un paso a otro, un valor fijo de f_{in} y f_{out} ocasiona que los índices de los errores tipo I y tipo II varíen. Algunos programas de computadora permiten que el analista especifique los niveles del error tipo I para f_{in} y f_{out}. Sin embargo, el nivel de significación "anunciado" en el programa no es el verdadero nivel, ya que la variable seleccionada es la que maximiza (o minimiza) el estadístico F parcial en ese paso. En ocasiones es útil experimentar con diferentes valores de f_{in} y f_{out} (o con diferentes índices del error tipo I anunciado) en varias corridas diferentes para ver si esto afecta de manera sustancial la elección del modelo final.

EJEMPLO 11-16

Se aplicará la regresión por pasos a los datos de la resistencia al desprendimiento del alambre adherido de la tabla 11-12. Quizás el lector quiera consultar la salida de computadora de la sección 11-11.3, la cual fundamenta los cálculos. En lugar de especificar los valores numéricos de f_{in} y f_{out}, se usa un error tipo I anunciado de $\alpha = 0.10$. El primer paso consiste en construir un modelo de regresión lineal simple usando la variable que produce el estadístico F más grande. Ésta es x_1, y como

$$f_1 = \frac{SS_R(\beta_1|\beta_0)}{MS_E(x_1)} = \frac{5\,885.8521}{9.5692} = 615.08 > f_{in} = f_{0.10,\,1,\,23} = 2.94$$

x_1 se incorpora en el modelo.

El segundo paso empieza encontrando la variable x_j que tiene el estadístico F parcial más grande, dado que x_1 está en el modelo. Ésta es x_4, y como

$$f_4 = \frac{SS_R(\beta_4|\beta_1, \beta_0)}{MS_E(x_4, x_1)} = \frac{116.4725}{4.7100} = 24.73 > f_{in} = f_{0.10,\,1,\,22} = 2.95$$

x_4 se incorpora en el modelo. Ahora el procedimiento evalúa si x_1 deberá conservarse o no, dado que x_4 está en el modelo. Esto implica calcular

$$f_1 = \frac{SS_R(\beta_1|\beta_4, \beta_0)}{MS_E(x_4, x_1)} = \frac{2\,582.5381}{4.7100} = 548.31 > f_{\text{out}} = f_{0.10,\,1,\,22} = 2.95$$

Por lo tanto, x_1 deberá conservarse. El paso 2 termina con ambas x_1 y x_4 en el modelo.

En el tercer paso se encuentra que la siguiente variable por incorporar es x_2. Puesto que

$$f_2 = \frac{SS_R(\beta_2|\beta_1, \beta_4, \beta_0)}{MS_E(x_2, x_4, x_1)} = \frac{40.2094}{3.0196} = 13.32 > f_{\text{in}} = f_{0.10,\,1,\,21} = 2.96$$

x_2 se incorpora en el modelo. Las pruebas F parciales de x_1 (dadas x_2 y x_4) y x_4 (dadas x_2 y x_1) indican que estas variables regresoras deberán conservarse. Por lo tanto, el tercer paso concluye con las variables x_1, x_2 y x_4 en el modelo.

En el cuarto paso, el procedimiento intenta agregar x_3. Sin embargo, el estadístico F parcial de x_3 es $f_3 = 0.59$, que es menor que $f_{\text{in}} = f_{0.10,\,1,\,20} = 2.97$. Por lo tanto, puesto que esta variable no puede agregarse y como ya no hay más variables regresoras candidatas por considerar, el procedimiento termina.

El procedimiento de regresión por pasos concluiría que el modelo de tres variables con x_1, x_2 y x_4 es la mejor ecuación de regresión. Sin embargo, deberán aplicarse a la ecuación las verificaciones estándares de la adecuación del modelo, tales como el análisis residual y la evaluación de gráficas C_p, a fin de asegurarse de que el modelo final es razonable, y de que ninguno de los supuestos de regresión se infringe seriamente. Obsérvese que este es el mismo modelo de regresión que se encontró con el método de todas las regresiones posibles.

Selección hacia adelante

El procedimiento de selección hacia adelante es una variante de la regresión por pasos y se basa en el principio de que las variables regresoras deberán agregarse al modelo una a la vez hasta que ya no queden variables regresoras candidatas que produzcan un incremento significativo en la suma de los cuadrados de la regresión. Es decir, las variables se agregan una a la vez siempre que su valor F parcial exceda f_{in}. La selección hacia adelante es una simplificación de la regresión por pasos que omite la prueba F parcial para eliminar las variables del modelo que se han agregado en pasos previos. Esta es una debilidad potencial de la selección hacia adelante; es decir, el procedimiento no explora el efecto que tiene agregar una variable regresora en el paso en curso sobre las variables regresoras agregadas en pasos anteriores.

EJEMPLO 11-17

La aplicación del algoritmo de selección hacia adelante a los datos de la resistencia al desprendimiento del alambre adherido de la tabla 11-12 empezaría incorporando x_1 en el modelo. Después, la variable regresora que produce la prueba F parcial más grande, dado que x_1 está en el modelo, se agrega: es la variable regresora x_4. En el tercer paso se incorpora x_2, la cual produce el estadístico F parcial más grande, dado que x_1 y x_4 están en el modelo. Puesto que el estadístico F parcial de x_3 no es significativo, el procedimiento termina. Obsérvese que en este ejemplo la selección hacia adelante lleva al mismo modelo final que la regresión por pasos. No siempre ocurrirá esto. En ocasiones, la selección hacia adelante produce un modelo diferente que la regresión por pasos.

Eliminación hacia atrás

Este algoritmo empieza considerando las K variables regresoras candidatas del modelo. Después se elimina la variable regresora con el estadístico F parcial menor, si este estadístico F no es significativo; es decir, si $f < f_{out}$. En seguida, se ajusta el modelo con $K-1$ variables regresoras, y se encuentra la siguiente variable regresora para su posible eliminación. El algoritmo termina cuando no puede eliminarse ninguna variable regresora.

EJEMPLO 11-18

Para aplicar la eliminación hacia atrás a los datos de la tabla 11-12, se empieza ajustando el modelo pleno o saturado con las cuatro variables regresoras. Este modelo es

$$\hat{y} = 1.06813 + 2.48423x_1 + 0.00855x_2 + 0.62337x_3 + 2.29392x_4$$

Las pruebas F parciales para cada variable son

$$f_1 = \frac{SS_R(\beta_1|\beta_2, \beta_3, \beta_4, \beta_0)}{MS_E} = 509.36$$

$$f_2 = \frac{SS_R(\beta_2|\beta_1, \beta_3, \beta_4, \beta_0)}{MS_E} = 13.15$$

$$f_3 = \frac{SS_R(\beta_3|\beta_1, \beta_2, \beta_4, \beta_0)}{MS_E} = 0.59$$

$$f_4 = \frac{SS_R(\beta_4|\beta_1, \beta_2, \beta_3, \beta_0)}{MS_E} = 9.39$$

Evidentemente, x_3 no contribuye al ajuste y deberá eliminarse. En el modelo de tres variables con variables regresoras (x_1, x_2, x_4), todas las variables regresoras son significativas por el criterio de la prueba F parcial y, por consiguiente, el algoritmo termina. Obsérvese que la eliminación hacia atrás ha resultado en el mismo modelo que se encontró con la selección hacia adelante y la regresión por pasos. Esto puede no ocurrir siempre.

Algunos comentarios sobre la selección del modelo final

Se han ilustrado varios enfoques diferentes para la selección de las variables en una regresión lineal múltiple. El modelo final obtenido con cualquier procedimiento para construir un modelo deberá someterse a las verificaciones de adecuación usuales, como el análisis residual, la prueba de falta de ajuste y el examen de los efectos de los puntos influyentes. El analista también puede considerar aumentar el conjunto original de variables candidatas con productos cruzados, términos polinomiales u otras transformaciones de las variables originales que podrían mejorar el modelo. Una crítica importante de los métodos de selección de variables tales como la regresión por pasos, es que el analista puede llegar a la conclusión de que existe una ecuación de regresión "mejor". En general, no es éste el caso, debido a que con frecuencia pueden usarse varios modelos de regresión igualmente buenos. Una manera de evitar este problema es usar varias técnicas de construcción de modelos y ver si se obtienen modelos diferentes. Por ejemplo, aquí se encontró el mismo modelo para los datos de la resistencia al desprendimiento del alambre

adherido usando la regresión por pasos, la selección hacia adelante y la eliminación hacia atrás. También se llegó al mismo modelo con el método de todas las regresiones posibles. Esta es una buena indicación de que el modelo de tres variables es la mejor ecuación de regresión.

Si el número de variables regresoras candidatas no es demasiado grande, se recomienda el método de todas las regresiones posibles. Los autores acostumbran recomendar los criterios de evaluación del MS_E y del C_p mínimos en conjunción con este procedimiento. Con el enfoque de todas las regresiones posibles puede encontrarse la "mejor" ecuación de regresión con respecto a estos criterios, en tanto que los métodos por pasos no ofrecen alguna garantía similar. Además, el procedimiento de todas las regresiones posibles no es distorsionado por las dependencias entre las variables regresoras, como es el caso de los métodos por pasos.

11-11.3 Salida de computadora de una regresión por pasos

Los programas de computadora son hoy parte indispensable de la construcción moderna de modelos de regresión, y se cuenta con varios programas excelentes. El programa Minitab llevará a cabo las regresiones por pasos básicas. Se ilustrará la salida de otro paquete de software ampliamente usado, el Statistical Analysis System (SAS). Se usarán los datos de la resistencia al desprendimiento del alambre adherido de la tabla 11-12.

En la figura 11-16 se presenta la salida de SAS PROC REG y PROC RSQUARE. PROC REG realizará la regresión por pasos, la selección hacia adelante, la eliminación hacia atrás y otras variantes del enfoque por pasos para la selección de variables. PROC RSQUARE es un algoritmo de todas las regresiones posibles. Obsérvese que los resultados obtenidos con este programa se usaron en los cálculos de la regresión por pasos ilustrados en la sección 11-11.2.

EJERCICIOS DE LA SECCIÓN 11-11

11-54. Use los datos del desempeño de los equipos de la Liga Nacional de Futbol Americano (NFL) del ejercicio 11-4 para construir modelos de regresión usando las siguientes técnicas:

a) Todas las regresiones posibles. Encuentre las ecuaciones que minimicen MS_E y que minimicen C_p.
b) Regresión por pasos.
c) Selección hacia adelante.
d) Eliminación hacia atrás.
e) Comente los varios modelos obtenidos. ¿Cuál de los modelos parece "mejor" y por qué?

11-55. Use los datos del rendimiento de gasolina del ejercicio 11-5 para construir modelos de regresión usando las siguientes técnicas:

a) Todas las regresiones posibles. Encuentre las ecuaciones de C_p mínimo y MS_E mínimo.
b) Regresión por pasos.
c) Selección hacia adelante.
d) Eliminación hacia atrás.
e) Comente los varios modelos obtenidos.

11-56. Considérense los datos del consumo de energía eléctrica del ejercicio 11-6. Construya modelos de regresión para los datos usando las siguientes técnicas:

a) Todas las regresiones posibles.
b) Regresión por pasos.
c) Selección hacia adelante.
d) Eliminación hacia atrás.
e) Comente los modelos obtenidos. ¿Cuál modelo se preferiría?

11-57. Considérense los datos de la resistencia al desprendimiento del alambre adherido del ejercicio 11-8. Construya modelos de regresión para los datos usando los siguientes métodos:

a) Regresión por pasos.
b) Selección hacia adelante.
c) Eliminación hacia atrás.

PROCEDIMIENTO DE SELECCIÓN HACIA ADELANTE PARA LA VARIABLE DEPENDIENTE Y

PASO 1 ENTRA LA VARIABLE X1 R CUADRADA = 0.93695437 C(P) = 50.46213232

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	1	5885.85206919	5885.85206919	615.08	0.0001
ERROR	23	220.09263481	9.56824499		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	5.11451557				
X1	2.90270442	0.11704072	5885.85206919	615.08	0.0001

COTAS DEL NÚMERO CONDICIÓN: 1, 1

PASO 2 ENTRA VARIABLE X4 R CUADRADA = 0.98302963 C(P) = 14.64453722

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	2	6002.32458234	3001.16229117	637.19	0.0001
ERROR	22	103.62012166	4.71000553		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	2.64573312				
X1	2.54772805	0.10880288	2582.53810667	548.31	0.0001
X4	3.54854459	0.71359082	116.47251315	24.73	0.0001

COTAS DEL NÚMERO CONDICIÓN: 1.7755752, 7.023007

PASO 3 ENTRA VARIABLE X2 R CUADRADA = 0.98961493 C(P) = 3.58888771

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	3	6042.53402205	2014.17800735	667.04	0.0001
ERROR	21	63.41068195	3.01955628		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	1.36706835				
X1	2.53491873	0.08718736	2552.49064020	815.32	0.0001
X2	0.00852152	0.00233520	40.20943970	13.32	0.0015
X4	2.59927756	0.62779136	51.76280083	17.14	0.0005

COTAS DEL NÚMERO CONDICIÓN: 2.119696, 15.86207

NINGUNA OTRA VARIABLE CUMPLE CON EL NIVEL DE SIGNIFICACIÓN DEL 0.1000 PARA ENTRAR EN EL MODELO.

Figura 11-16 Salida de los programas SAS PROC REG y PROC RSQUARE para los datos de la resistencia al desprendimiento del alambre adherido.

RESUMEN DEL PROCEDIMIENTO DE SELECCIÓN HACIA ADELANTE
PARA LA VARIABLE DEPENDIENTE Y

PASO	VARIABLE INCORPORADA	NÚMERO PRESENTE	R**2 PARCIAL	R**2 DEL MODELO	C(P)	F	PROB>F
1	X1	1	0.9640	0.9640	50.4621	615.0801	0.0001
2	X4	2	0.0191	0.9830	14.6445	24.7287	0.0001
3	X2	3	0.0066	0.9896	3.5889	13.3163	0.0015

PROCEDIMIENTO DE ELIMINACIÓN HACIA ATRÁS PARA LA VARIABLE DEPENDIENTE Y

PASO 0 TODAS LAS VARIABLES SE INCORPORAN R CUADRADA = 0.98991196 C(P)= 5.00000000

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	4	6044.34770775	1511.08692694	490.64	0.0001
ERROR	20	61.59699625	3.07984981		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	1.06813405				
X1	2.48423244	0.11007295	1568.74664163	509.36	0.0001
X2	0.00855369	0.00235877	40.50082723	13.15	0.0017
X3	0.62236914	0.81102045	1.81368571	0.59	0.4518
X4	2.29392355	0.74854920	28.92322543	9.39	0.0061

COTAS DEL NÚMERO CONDICIÓN: 3.980888, 44.3724

PASO 1 SE ELIMINA LA VARIABLE X3 R CUADRADA = 0.98961493 C(P)= 3.58888771

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	3	6042.53402205	2014.17800735	667.04	0.0001
ERROR	21	63.41068195	3.01955628		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	1.36706835				
X1	2.53491873	0.08718736	2552.49064020	845.32	0.0001
X2	0.00852152	0.00233520	40.20943970	13.32	0.0015
X4	2.59927756	0.62779136	51.76280083	17.14	0.0005

COTAS DEL NÚMERO CONDICIÓN: 2.119696, 15.86207

TODAS LAS VARIABLES DEL MODELO SON SIGNIFICATIVAS EN EL NIVEL 0.1000.

RESUMEN DEL PROCEDIMIENTO DE ELIMINACIÓN HACIA ATRÁS
PARA LA VARIABLE DEPENDIENTE Y

PASO	VARIABLE ELIMINADA	NÚMERO PRESENTE	R**2 PARCIAL	R**2 DEL MODELO	C(P)	F	PROB>F
1	X1	3	0.0003	0.9896	3.5889	0.5889	0.4518

Figura 11-16 (*Continuación.*)

PROCEDIMIENTO DE REGRESIÓN POR PARTES PARA LA VARIABLE DEPENDIENTE Y

NOTA: SLSTAY SE HA FIJADO EN .15 PARA LA TÉCNICA POR PASOS.

PASO 1 ENTRA LA VARIABLE X1 R CUADRADA = 0.96395437 C(P) = 50.46213232

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	1	5885.85206919	5885.85206919	615.08	0.0001
ERROR	23	220.09263481	9.56824499		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	5.11451557				
X1	2.90270442	0.11704072	5885.85206919	615.08	0.0001

COTAS DEL NÚMERO CONDICIÓN: 1, 1

PASO 2 ENTRA VARIABLE X4 R CUADRADA = 0.98302963 C(P) = 14.64453722

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	2	6002.32458234	3001.16229117	637.19	0.0001
ERROR	22	103.62012166	4.71000553		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	2.64573312				
X1	2.54772805	0.10880288	2582.53810667	548.31	0.0001
X4	3.54854459	0.71359082	116.47251315	24.73	0.0001

COTAS DEL NÚMERO CONDICIÓN: 1.7755752, 7.023007

PASO 3 ENTRA VARIABLE X2 R CUADRADA = 0.98961493 C(P) =3.58888771

	GL	SUMA DE CUADRADOS	CUADRADO MEDIO	F	PROB>F
REGRESIÓN	3	6042.53402205	2014.17800735	667.04	0.0001
ERROR	21	63.41068195	3.01955628		
TOTAL	24	6105.94470400			

	VALOR B	ERROR ESTÁNDAR	SS TIPO II	F	PROB>F
ORDENADA AL ORIGEN	1.36706835				
X1	2.53491873	0.08718736	2552.49064020	845.32	0.0001
X2	0.00852152	0.00233520	40.20943970	13.32	0.0015
X4	2.59927756	0.62779136	51.76280083	17.14	0.0005

COTAS DEL NÚMERO CONDICIÓN: 2.119696, 15.86207

NINGUNA OTRA VARIABLE CUMPLE CON EL NIVEL DE SIGNIFICACIÓN DEL 0.1000 PARA ENTRAR EN EL MODELO.

Figura 11-16 (*Continuación.*)

RESUMEN DEL PROCEDIMIENTO DE REGRECIÓN PARA LA VARIABLE DEPENDIENTE Y

PASO	VARIABLE INCORPORADA	VARIABLE ELIMINADA	NÚMERO PRESENTE	R**2 PARCIAL	R**2 DEL MODELO	C(P)	F	PROB>F
1	X1		1	0.9640	0.9640	50.4621	615.0801	0.0001
2	X4		2	0.0191	0.9830	14.6445	24.7287	0.0001
3	X2		3	0.0066	0.9896	3.5889	13.3163	0.0015

N = 25 MODELOS DE REGRESIÓN PARA LA VARIABLE DEPENDIENTE: MODELO Y:

MODELO 1

NÚMERO PRESENTE EN EL MODELO	R CUADRADA	C(P)	VARIABLES EN EL MODELO			
1	0.24291747	1479.951	X2			
1	0.56007492	851.172	X4			
1	0.69830969	577.115	X3			
1	0.96395437	50.462132	X1			
2	0.57158123	830.360	X2	X4		
2	0.72162460	532.892	X2	X3		
2	0.72232683	531.500	X3	X4		
2	0.97220237	36.110087	X1	X3		
2	0.98113748	18.395811	X1	X2		
2	0.98302963	14.644537	X1	X4		
3	0.73299076	512.358	X2	X3	X4	
3	0.98327895	16.150260	X1	X3	X4	
3	0.98517507	12.391116	X1	X2	X3	
3	0.98961493	3.588888	X1	X2	X4	
4	0.98991196	5.000000	X1	X2	X3	X4

Figura 11-16 (*Continuación.*)

d) Comente los modelos obtenidos. ¿Cuál modelo se preferiría?

11-58. Considérense los datos de la NHL del ejercicio 11-11. Construya modelos de regresión para estos datos usando los siguientes métodos:

a) Regresión por pasos.
b) Selección hacia adelante.
c) Eliminación hacia atrás.
d) ¿Cuál modelo se preferiría?

11-59. Considérense los datos del ejercicio 11-51. Use todos los términos del modelo cuadrático pleno o saturado como variables regresoras candidatas.

a) Use la selección hacia adelante para identificar un modelo.
b) Use la eliminación hacia atrás para identificar un modelo.
c) Compare los dos modelos obtenidos en los incisos a) y b). ¿Cuál modelo se preferiría y por qué?

11-60. Encuentre la ecuación del C_p mínimo y la ecuación que maximice el estadístico R^2 ajustado, para los datos de la resistencia al desprendimiento del alambre adherido del ejercicio 11-8. ¿La misma ecuación satisface ambos criterios?

11-61. Para los datos de la NHL del ejercicio 11-11,

a) Encuentre la ecuación que minimice C_p.

b) Encuetre la ecuación que minimice MS_E.

c) Encuentre la ecuación que maximice el R^2 ajustado. ¿Es la misma ecuación que se encontró en el inciso b)?

11-62. Se ha usado una muestra de 30 observaciones para ajustar un modelo de regresión. El modelo pleno o saturado tiene nueve variables regresoras, la estimación de la varianza es $\sigma^2 = MS_E = 100$, y $R^2 = 0.92$.

a) Calcule el estadístico F para probar la significación de la regresión. Usando $\alpha = 0.05$, ¿a qué conclusión se llegaría?

b) Suponga que se ajusta otro modelo usando sólo cuatro de las variables regresoras originales y que la suma de los cuadrados del error para este nuevo modelo es 2 200. Encuentre la estimación de σ^2 para este nuevo modelo reducido. ¿Se concluiría que el modelo reducido es superior al original? ¿Por qué?

c) Encuentre el valor de C_p para el modelo reducido del inciso b). ¿Se concluiría que el modelo reducido es mejor que el modelo original?

11-63. Se usa una muestra de 25 observaciones para ajustar un modelo de regresión con siete variables. La estimación de σ^2 para este modelo pleno o saturado es $MS_E = 10$.

a) Un algoritmo de selección hacia adelante ha incluido tres de las siete variables regresoras originales en el modelo. La suma de los cuadrados del error del modelo de tres variables es $SS_E = 300$. Con base en C_p, ¿se concluiría que en el modelo de tres variables aún está presente algún sesgo?

b) Después de considerar el modelo de selección hacia adelante del inciso a), suponga que quiere agregarse una variable regresora más al modelo. Esta variable regresora reducirá la suma de los cuadrados del error a 275. ¿La incorporación de esta variable mejora el modelo? ¿Por qué?

11-12 COLINEALIDAD MÚLTIPLE

En los problemas de regresión múltiple se espera encontrar dependencias entre la variable de respuesta Y y las variables regresoras x_j. Sin embargo, en la mayoría de los problemas de regresión se encuentra que también hay dependencias entre las variables regresoras x_j. Cuando estas dependencias son fuertes, se dice que existe **colinealidad múltiple**. La colinealidad múltiple puede tener efectos de consecuencia sobre las estimaciones de los coeficientes de regresión y sobre la aplicabilidad general del modelo estimado.

Los efectos de colinealidad múltiple pueden demostrarse con facilidad. Los elementos de la diagonal de la matriz $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1}$ pueden escribirse como

$$C_{jj} = \frac{1}{\left(1 - R_j^2\right)} \qquad j = 1, 2, \ldots, k$$

donde R_j^2 es el coeficiente de determinación múltiple que resulta de hacer la regresión de x_j sobre las $k - 1$ variables regresoras restantes. Evidentemente, entre más fuerte sea la dependencia lineal de x_j sobre las variables regresoras restantes, y por tanto entre más fuerte sea la colinealidad múltiple, más grande será el valor de R_j^2. Recuérdese que $V(\hat{\beta}_j) = \sigma^2 C_{jj}$. Por lo tanto, se dice que la varianza de $\hat{\beta}_j$ está "inflada" por la cantidad $(1 - R_j^2)^{-1}$. Por consiguiente, el **factor de inflación de la varianza** de $\hat{\beta}_j$ se define como

$$VIF(\hat{\beta}_j) = \frac{1}{(1-R_j^2)} \qquad j = 1, 2, \ldots, k \tag{11-50}$$

Estos factores son una medida importante del grado en que está presente la colinealidad múltiple.

Aun cuando las estimaciones de los coeficientes de regresión son muy imprecisas cuando está presente la colinealidad múltiple, la ecuación del modelo ajustado puede seguir siendo de utilidad. Por ejemplo, suponga que quieren predecirse nuevas observaciones de la variable de respuesta. Si estas predicciones son interpolaciones en la región original del espacio x donde está presente la colinealidad múltiple, entonces la mayoría de las veces se obtendrán predicciones satisfactorias, debido a que aun cuando las estimaciones de las β_j individuales pueden ser pobres, la función $\sum_{j=1}^{k} \beta_j x_{ij}$ puede estimarse bastante bien. Por otra parte, si la predicción de nuevas observaciones requiere hacer una extrapolación fuera de la región original del espacio x donde se colectaron los datos, entonces por lo general se esperaría obtener resultados pobres. Una extrapolación requiere normalmente buenas estimaciones de los parámetros individuales del modelo.

La colinealidad múltiple surge por varias razones. Ocurrirá cuando el analista recaba datos tales que una restricción lineal se encuentra presente aproximadamente entre las columnas de la matriz **X**. Por ejemplo, si cuatro variables regresoras son los componentes de una mezcla, entonces esta restricción existirá siempre, porque la suma de los componentes es siempre constante. Por lo general, estas restricciones no se dan de manera precisa, y el analista podría ignorar su existencia.

Hay varias formas de detectar la presencia de la colinealidad múltiple. Dos de las más fáciles de entender se analizarán brevemente.

1. Los **factores de inflación de la varianza**, definidos en la ecuación 11-50, son medidas de colinealidad múltiple muy útiles. Entre mayor sea el factor de inflación de la varianza, más marcada será la colinealidad múltiple. Algunos autores han sugerido que si cualquiera de los factores de inflación de la varianza excede 10, entonces la colinealidad múltiple constituye un problema. Otros autores consideran este valor muy dadivoso y sugieren que los factores de inflación de la varianza no deberán exceder 4 o 5. El programa Minitab calcula los factores de inflación de la varianza. En la tabla 11-4 se presenta la salida de la regresión múltiple de Minitab para los datos de la resistencia al desprendimiento del alambre adherido. Puesto que tanto VIF_1 como VIF_2 son pequeños, no hay ningún problema con la colinealidad múltiple.

2. Si la prueba F de la significación de la regresión es significativa, pero las pruebas de los coeficientes de regresión individuales son no significativas, entonces puede estar presente la colinealidad múltiple.

Se han propuesto varias medidas correctivas para resolver el problema de la colinealidad múltiple. Se sugiere con frecuencia aumentar los datos con nuevas observaciones diseñadas específicamente para romper las dependencias lineales aproximadas que existen normalmente. Sin embargo, en ocasiones esto es imposible debido a razones económicas o por las restricciones físicas que relacionan las x_j. Otra posibilidad es eliminar ciertas variables del modelo, pero este enfoque tiene la desventaja de descartar la información contenida en las variables eliminadas.

Puesto que la colinealidad múltiple afecta principalmente la estabilidad de los coeficientes de regresión, al parecer estimar estos parámetros con algún método que sea menos sensitivo a la colinealidad múltiple que los mínimos cuadrados ordinarios, sería de ayuda. Se han sugerido varios métodos. Una alternativa de los mínimos cuadrados ordinarios, la **regresión sobre ascensos** (cordilleras), puede ser útil para combatir la colinealidad múltiple. Para más detalles sobre la regresión sobre ascensos, véase Montgomery y Peck y Myers.

Ejercicios complementarios

11-64. Los datos que se presentan en la tabla de la página 557 representan el empuje de un motor de turbina de avión (y) y seis variables regresoras candidatas x_1 = velocidad de rotación primaria, x_2 = velocidad de rotación secundaria, x_3 = rapidez del flujo del combustible, x_4 = presión, x_5 = temperatura de los gases de escape, y x_6 = temperatura ambiente en el momento de la prueba.

a) Ajuste un modelo de regresión lineal múltiple usando x_3 = rapidez del flujo del combustible, x_4 = presión y x_5 = temperatura de los gases de escape como variables regresoras.

b) Pruebe la significación de la regresión usando $\alpha = 0.01$. Encuentre el valor P para esta prueba. ¿A qué conclusiones se llega?

c) Encuentre el estadístico de la prueba t para cada variable regresora. Usando $\alpha = 0.01$, explique cuidadosamente la conclusión que puede sacarse a partir de estos estadísticos.

d) Encuentre R^2 y el estadístico ajustado para este modelo.

e) Construya una gráfica de probabilidad normal de los residuales e interprétela.

f) Grafique los residuales contra $\hat{y}$. ¿Hay alguna indicación de la desigualdad de la varianza o de no linealidad?

g) Grafique los residuales contra x_3. ¿Hay alguna indicación de no linealidad?

h) Prediga el empuje de un motor para el que $x_3 = 1\,670$, $x_4 = 170$ y $x_5 = 1\,589$.

11-65. Considérense los datos del empuje del motor del ejercicio 11-64. Ajuste de nuevo el modelo usando $y^* = ln\ y$ como variable de respuesta y $x_3^* = ln\ x_3$ como variable regresora (junto con x_4 y x_5).

a) Pruebe la significación de la regresión usando $\alpha = 0.01$. Encuentre el valor P para esta prueba y establezca sus conclusiones.

b) Use el estadístico t para probar H_0: $\beta_j = 0$ contra H_1: $\beta_j \neq 0$ para cada una de las variables del modelo. Si $\alpha = 0.01$, ¿qué conclusiones pueden sacarse?

c) Grafique los residuales contra $\hat{y}^*$ y contra x_3^*. Comente estas gráficas. ¿Cuál es el resultado de compararlas con sus contrapartes obtenidas en los incisos f) y g) del ejercicio 11-64?

11-66. Los puntos transitorios de un convertidor electrónico están influidos por muchos factores. En la página 558 se presentan los datos sobre el punto transitorio (y, en volts) de convertidores PMOS-NMOS y cinco variables regresoras candidatas: x_1 = anchura del dispositivo NMOS, x_2 = largo del dispositivo NMOS, x_3 = anchura del dispositivo PMOS, x_4 = largo del dispositivo PMOS, y x_5 = temperatura (°C).

a) Ajuste el modelo de regresión lineal múltiple a estos datos. Pruebe la significación de la regresión usando $\alpha = 0.01$. Encuentre el valor P para esta prueba y úselo para sacar conclusiones.

b) Pruebe la contribución de cada variable al modelo usando la prueba t con $\alpha = 0.05$. ¿Cuáles son las conclusiones?

c) Elimine x_5 del modelo. Pruebe el nuevo modelo para la significación de la regresión. Pruebe también la contribución relativa de cada variable regresora al nuevo modelo con la prueba t. Usando $\alpha = 0.05$, ¿cuáles son las conclusiones?

Número de observación	y	x_1	x_2	x_3	x_4	x_5	x_6
1	4 540	2 140	20 640	30 250	205	1 732	99
2	4 315	2 016	20 280	30 010	195	1 697	100
3	4 095	1 905	19 860	29 780	184	1 662	97
4	3 650	1 675	18 980	29 330	164	1 598	97
5	3 200	1 474	18 100	28 960	144	1 541	97
6	4 833	2 239	20 740	30 083	216	1 709	87
7	4 617	2 120	20 305	29 831	206	1 669	87
8	4 340	1 990	19 961	29 604	196	1 640	87
9	3 820	1 702	18 916	29 088	171	1 572	85
10	3 368	1 487	18 012	28 675	149	1 522	85
11	4 445	2 107	20 520	30 120	195	1 740	101
12	4 188	1 973	20 130	29 920	190	1 711	100
13	3 981	1 864	19 780	29 720	180	1 682	100
14	3 622	1 674	19 020	29 370	161	1 630	100
15	3 125	1 440	18 030	28 940	139	1 572	101
16	4 560	2 165	20 680	30 160	208	1 704	98
17	4 340	2 048	20 340	29 960	199	1 679	96
18	4 115	1 916	19 860	29 710	187	1 642	94
19	3 630	1 658	18 950	29 250	164	1 576	94
20	3 210	1 489	18 700	28 890	145	1 528	94
21	4 330	2 062	20 500	30 190	193	1 748	101
22	4 119	1 929	20 050	29 960	183	1 713	100
23	3 891	1 815	19 680	29 770	173	1 684	100
24	3 467	1 595	18 890	29 360	153	1 624	99
25	3 045	1 400	17 870	28 960	134	1 569	100
26	4 411	2 047	20 540	30 160	193	1 746	99
27	4 203	1 935	20 160	29 940	184	1 714	99
28	3 968	1 807	19 750	29 760	173	1 679	99
29	3 531	1 591	18 890	29 350	153	1 621	99
30	3 074	1 388	17 870	28 910	133	1 561	99
31	4 350	2 071	20 460	30 180	198	1 729	102
32	4 128	1 944	20 010	29 940	186	1 692	101
33	3 940	1 831	19 640	29 750	178	1 667	101
34	3 480	1 612	18 710	29 360	156	1 609	101
35	3 064	1 410	17 780	28 900	136	1 552	101
36	4 402	2 066	20 520	30 170	197	1 758	100
37	4 180	1 954	20 150	29 950	188	1 729	99
38	3 973	1 835	19 750	29 740	178	1 690	99
39	3 530	1 616	18 850	29 320	156	1 616	99
40	3 080	1 407	17 910	28 910	137	1 569	100

Número de observación	x_1	x_2	x_3	x_4	x_5	y
1	3	3	3	3	0	0.787
2	8	30	8	8	0	0.293
3	3	6	6	6	0	1.710
4	4	4	4	12	0	0.203
5	8	7	6	5	0	0.806
6	10	20	5	5	0	4.713
7	8	6	3	3	25	0.607
8	6	24	4	4	25	9.107
9	4	10	12	4	25	9.210
10	16	12	8	4	25	1.365
11	3	10	8	8	25	4.554
12	8	3	3	3	25	0.293
13	3	6	3	3	50	2.252
14	3	8	8	3	50	9.167
15	4	8	4	8	50	0.694
16	5	2	2	2	50	0.379
17	2	2	2	3	50	0.485
18	10	15	3	3	50	3.345
19	15	6	2	3	50	0.208
20	15	6	2	3	75	0.201
21	10	4	3	3	75	0.329
22	3	8	2	2	75	4.966
23	6	6	6	4	75	1.362
24	2	3	8	6	75	1.515
25	3	3	8	8	75	0.751

d) Obsérvese que el MS_E del modelo del inciso c) es menor que el MS_E del modelo pleno o saturado del inciso a). Explique por qué ocurre esto.

e) Calcule los residuales studentizados. ¿Alguno de ellos parece ser excepcionalmente grande?

f) Suponga que el analista se enteró de que la segunda observación se registró incorrectamente. Elimine esta observación y ajuste de nuevo el modelo usando x_1, x_2, x_3 y x_4 como variables regresoras. Obsérvese que R^2 de este modelo es considerablemente más grande que R^2 de cualquiera de los modelos ajustados anteriormente. Explique por qué se ha incrementado R^2 en este modelo.

g) Pruebe el modelo del inciso f) para la significación de la regresión usando $\alpha = 0.05$. Investigue asimismo la contribución de cada variable regresora al modelo usando la prueba t con $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?

h) Grafique los residuales del modelo del inciso f) contra $\hat{y}$ y contra cada una de las variables regresoras x_1, x_2, x_3 y x_4. Comente las gráficas.

11-67. Considérense los datos del convertidor del ejercicio 11-66. Elimine la observación 2 de los datos originales. Defina las nuevas variables como sigue: $y^* = ln\, y$, $x_1^* = 1/\sqrt{x_1}$, $x_2^* = \sqrt{x_2}$, $x_3^* = 1/\sqrt{x_3}$, y $x_4^* = \sqrt{x_4}$.

a) Ajuste un modelo de regresión usando estas variables regresoras transformadas (no use x_5).

b) Pruebe el modelo para la significación de la regresión usando $\alpha = 0.05$. Use la prueba t para investigar la contribución de cada variable al modelo ($\alpha = 0.05$). ¿Cuáles son sus conclusiones?

c) Grafique los residuales contra $\hat{y}^*$ y contra cada una de las variables regresoras transformadas. Comente las gráficas.

11-68. Se presentan a continuación los datos para y = licor verde (g/l) y x = velocidad de la máquina de papel (pies/min) de una máquina de papel Kraft. (Los datos se tomaron de una gráfica en un artículo de *Journal Tappi.*)

y	16.0	15.8	15.6	15.5	14.8
x	1 700	1 720	1 730	1 740	1 750
y	14.0	13.5	13.0	12.0	11.0
x	1 760	1 770	1 780	1 790	1 795

a) Ajuste el modelo $Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$ usando mínimos cuadrados.

b) Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Cuáles son sus conclusiones?

c) Pruebe la contribución del término cuadrático al modelo, sobre la contribución del término lineal, usando un estadístico F. Si $\alpha = 0.05$, ¿qué conclusiones pueden sacarse?

d) Grafique los residuales del modelo del inciso a) contra $\hat{y}$. ¿La gráfica revela alguna falta de adecuación?

e) Construya una gráfica de probabilidad normal de los residuales. Comente el supuesto de normalidad.

11-69. Considérense los datos del empuje de la turbina del ejercicio 11-64.

a) Use el método de todas las regresiones posibles para seleccionar la mejor ecuación de regresión, donde el modelo con el valor mínimo de MS_E se seleccionará como el "mejor".

b) Repita el inciso a) usando el criterio C_p para identificar la mejor ecuación.

c) Use la regresión por pasos para seleccionar un modelo de regresión de un subconjunto de variables.

d) Compare los modelos obtenidos en los incisos a), b) y c).

11-70. Considérense los datos del convertidor electrónico de los ejercicios 11-66 y 11-67. Defina las variables de respuesta y de regresión como en el ejercicio 11-67, y elimine de la muestra la segunda observación.

a) Use la técnica de todas las regresiones posibles para encontrar la ecuación que minimice C_p.

b) Use todas las regresiones posibles para encontrar la ecuación que minimice MS_E.

c) Use la regresión por pasos para seleccionar un modelo de regresión de un subconjunto de variables.

d) Compare los modelos obtenidos.

11-71. Considérese el modelo de regresión de tres variables para los datos del empuje de la turbina del ejercicio 11-65. Calcule los factores de inflación de la varianza para este modelo. ¿Se concluiría que la colinealidad múltiple es un problema en este modelo?

11-72. Se usó un modelo de regresión múltiple para relacionar y = viscosidad de un producto químico con x_1 = temperatura y x_2 = tiempo de reacción. El conjunto de datos constó de $n = 15$ observaciones.

a) Los coeficientes de regresión estimados fueron $\hat{\beta}_0 = 300.00$, $\hat{\beta}_1 = 0.85$ y $\hat{\beta}_2 = 10.40$. Calcule una estimación de la viscosidad media cuando $x_1 = 100°F$ y $x_2 = 2$ hr.

b) Las sumas de los cuadrados fueron $SS_T = 1\,230.50$ y $SS_E = 120.30$. Pruebe la significación de la regresión usando $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse?

c) ¿Qué proporción de la variabilidad total de la viscosidad es explicada por las variables del modelo?

d) Suponga que se agrega otra variable regresora al modelo, x_3 = velocidad del agitador. El nuevo valor de la suma de los cuadrados del error es $SS_E = 117.20$. ¿La incorporación de la nueva variable tiene como resultado un valor menor de MS_E? Comente la significación de este resultado.

e) Calcule un estadístico F para determinar la contribución de x_3 al modelo. Usando $\alpha = 0.05$, ¿a qué conclusiones puede llegarse?

11-73. Un artículo de *Journal of the American Ceramics Society* (Vol. 75, pp. 112-116) describe un proceso para inmovilizar desechos químicos o nucleares en el suelo disolviendo el suelo contaminado en un bloque de vidrio. Los autores mezclan CaO y Na_2O con suelo y modelan la viscosidad y la conductividad eléctrica. El modelo de la conductividad eléctrica incluye seis variables regresoras, y la muestra consta de $n = 14$ observaciones.

a) Para el modelo de seis variables regresoras, suponer que $SS_T = 0.50$ y $R^2 = 0.94$. Encuentre SS_E y SS_R, y use esta información para probar la significación de la regresión con $\alpha = 0.05$. ¿Cuáles son sus conclusiones?

b) Suponga que se elimina del modelo una de las variables regresoras originales, lo cual resulta en $R^2 = 0.92$. ¿Qué puede concluirse acerca de la contribución de la variable que se eliminó? Responda esta pregunta calculando un estadístico F.

c) ¿La eliminación de la variable regresora del inciso b) da como resultado un valor menor de MS_E para el modelo de cinco variables, en comparación con el modelo original de seis variables? Comente la significación de su respuesta.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

11-74. Considérese un modelo de regresión múltiple con k variables regresoras. Demuestre que el estadístico de prueba para la significación de la regresión puede escribirse como

$$F_0 = \frac{R^2/k}{(1-R^2)/(n-k-1)}$$

Suponga que $n = 20$, $k = 4$ y $R^2 = 0.90$. Si $\alpha = 0.05$, ¿qué conclusiones se sacarían acerca de la relación entre y y las cuatro variable regresoras?

11-75. Se usa un modelo de regresión para relacionar una variable de respuesta y con $k = 4$ variables regresoras con $n = 20$. ¿Cuál es el valor menor de R^2 que dará como resultado una regresión significativa si $\alpha = 0.05$? Use los resultados del ejercicio anterior. ¿Resulta sorprendente lo pequeño que es el valor de R^2?

11-76. Demuestre que los residuales de un modelo de regresión múltiple pueden expresarse como $\boldsymbol{e} = (\mathbf{I} - \mathbf{H})\mathbf{y}$, donde $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

11-77. Demuestre que la varianza del residuo i-ésimo, e_i, en un modelo de regresión múltiple es $\sigma(1 - h_{ii})$ y que la covarianza entre e_i y e_j es $-\sigma^2 h_{ij}$, donde las h son los elementos de $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

11-78. Considérese el modelo de regresión lineal múltiple $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Si $\hat{\boldsymbol{\beta}}$ denota el estimador de mínimos cuadrados de $\boldsymbol{\beta}$, demuestre que $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \mathbf{R}\boldsymbol{\varepsilon}$, donde $\mathbf{R} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

11-79. **Mínimos cuadrados restringidos.** Suponga que quiere encontrarse el estimador de mínimos cuadrados de $\boldsymbol{\beta}$ en el modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ sujeto a un conjunto de restricciones de igualdad, digamos, $\mathbf{T}\boldsymbol{\beta} = \mathbf{c}$.

a) Demuestre que el estimador es

$$\hat{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1} \times \mathbf{T}'\left[\mathbf{T}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}'\right]^{-1}\left(\mathbf{c} - \mathbf{T}\hat{\boldsymbol{\beta}}\right)$$

donde $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

b) Comente las situaciones en las que este modelo podría ser apropiado.

11-80. **Regresión lineal por partes (I).** Suponga que y tiene una relación lineal por partes con x. Es decir, son apropiadas diferentes relaciones lineales en los intervalos $-\infty < x \leq x^*$ y $x^* < x < \infty$. Establezca cómo pueden usarse variables indicadoras para ajustar un modelo de regresión lineal por partes como éste, suponiendo que el punto x^* es conocido.

11-81. **Regresión lineal por partes (II).** Considérese el modelo de regresión lineal por partes descrito en el ejercicio 11-80. Suponga que en el punto x^* ocurre una discontinuidad en la función de regresión. Establezca cómo pueden usarse variables indicadoras para incorporar la discontinuidad en el modelo.

11-82. **Regresión lineal por partes (III).** Considérese el modelo de regresión lineal por partes descrito en el ejercicio 11-80. Suponga que no se conoce el punto x^* con certeza y debe estimarse. Sugiera un enfoque que pueda usarse para ajustar el modelo de regresión lineal por partes.

12 Diseño y análisis de experimentos con un solo factor: el análisis de varianza

PLAN GENERAL DEL CAPÍTULO

12-1 ESTRATEGIA DE EXPERIMENTACIÓN

12-2 EXPERIMENTO CON UN SOLO FACTOR COMPLETAMENTE ALEATORIZADO

- 12-2.1 Un ejemplo
- 12-2.2 El análisis de varianza
- 12-2.3 Salida de computadora
- 12-2.4 Análisis de residuales y verificación del modelo

12-3 PRUEBAS SOBRE MEDIAS DE TRATAMIENTOS INDIVIDUALES

- 12-3.1 Comparación gráfica de medias
- 12-3.2 Contrastes ortogonales
- 12-3.3 Método de la diferencia significativa mínima de Fisher

12-4 MODELO DE EFECTOS ALEATORIOS

12-5 DISEÑO DE BLOQUES COMPLETOS ALEATORIZADOS

- 12-5.1 Diseño y análisis estadístico
- 12-5.2 Pruebas sobre medias de tratamientos individuales
- 12-5.3 Análisis de residuales y verificación del modelo
- 12-5.4 Diseño de bloques completos aleatorizados con factores aleatorios

12-6 DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA EN EXPERIMENTOS CON UN SOLO FACTOR

- 12-6.1 Caso con efectos fijos
- 12-6.2 Caso con efectos aleatorios

12-1 ESTRATEGIA DE EXPERIMENTACIÓN

Los experimentos son parte natural en el proceso de la toma de decisiones en la ingeniería y las ciencias. Supóngase, por ejemplo, que un ingeniero civil investiga los efectos de diferentes métodos de curado sobre la resistencia a la compresión media del concreto. El experimento consistiría en formar varios ejemplares de prueba de concreto utilizando cada uno de los métodos de curado propuestos para después probar la resistencia a la compresión de cada ejemplar de prueba. Los datos de este experimento podrían usarse a fin de determinar qué método de curado deberá usarse para conseguir la resistencia a la compresión media máxima.

Si sólo hay dos métodos de curado que sean de interés, este experimento podría diseñarse y analizarse aplicando los métodos para hipótesis estadísticas con dos muestras introducidos en el capítulo 9. Es decir, el experimentador tiene un solo **factor** de interés (el método de curado) y sólo hay dos **niveles** del factor. Si el experimentador quiere determinar cuál de los métodos de curado produce la resistencia a la compresión máxima, entonces puede determinar el número de ejemplares por probar con las curvas de operación característica de la carta VI del apéndice, y aplicar la prueba t para decidir si las dos medias difieren.

En muchos experimentos con un solo factor, es necesario considerar más de dos niveles del factor. Por ejemplo, quizás el ingeniero civil quiera investigar cinco métodos de curado diferentes. En este capítulo se explica cómo puede usarse el **análisis de varianza** para comparar medias cuando hay más de dos niveles de un solo factor. Se revisará asimismo la **aleatorización** de las corridas experimentales y el importante papel que desempeña este concepto en la estrategia global de experimentación. En el capítulo siguiente se explicará cómo diseñar y analizar experimentos con varios factores.

Las técnicas de diseño experimental basadas en la estadística son particularmente útiles en el campo de la ingeniería para mejorar el desempeño de un proceso de manufactura. Tienen también diversas aplicaciones en el desarrollo de nuevos procesos. La mayoría de los procesos pueden describirse en términos de varias **variables controlables**, como la temperatura, la presión y la rapidez de alimentación. Mediante el uso de experimentos diseñados, los ingenieros pueden determinar el subconjunto de las variables del proceso que ejerce la mayor influencia sobre el desempeño del proceso. Los resultados de un experimento como éste pueden llevar a

1. El mejoramiento del rendimiento del proceso.
2. La reducción de la variabilidad del proceso y a una concordancia más próxima con los requerimientos nominales o proyectados.
3. La reducción del tiempo de diseño y desarrollo.
4. La reducción en el costo de operación.

Los métodos del diseño experimental también son útiles en las actividades de **diseño de ingeniería**, donde se desarrollan nuevos productos y se mejoran los existentes. Algunas aplicaciones típicas de los experimentos diseñados estadísticamente en el diseño de ingeniería incluyen

1. Evaluación y comparación de configuraciones de diseño básico.
2. Evaluación de diferentes materiales.

3. Selección de parámetros de diseño para que el producto tenga un buen funcionamiento en una amplia variedad de condiciones de campo (o para que el diseño sea robusto).
4. Determinación de parámetros de diseño clave de un producto, que inciden en el desempeño del mismo.

La utilización del diseño experimental en el proceso de ingeniería de diseño puede redundar en productos cuya fabricación sea más sencilla, en productos con un mejor desempeño de campo y confiabilidad que sus competidores, y en productos que pueden diseñarse, desarrollarse y producirse en menos tiempo.

Los experimentos diseñados generalmente se realizan **secuencialmente**. Es decir, el primer experimento con un sistema complejo (quizás un proceso de manufactura) que incluye muchas variables controlables es con frecuencia un **experimento exhaustivo de depuración** diseñado para determinar cuáles son las variables más importantes. Se realizan experimentos subsecuentes para refinar esta información y determinar los ajustes necesarios en estas variables críticas para mejorar el proceso. Por último, el objetivo del experimentador es la optimización, es decir, determinar los niveles de las variables críticas que producen el desempeño óptimo del proceso.

Todo experimento incluye una secuencia de actividades:

1. **Conjetura**: la hipótesis original que motiva el experimento.
2. **Experimento**: la prueba realizada para investigar la conjetura.
3. **Análisis**: el análisis estadístico de los datos del experimento.
4. **Conclusión**: qué se ha aprendido acerca de la conjetura original del experimento. Con frecuencia el experimento llevará a una revisión de la conjetura, y a un nuevo experimento, y así sucesivamente.

Los métodos estadísticos introducidos en este capítulo y el capítulo 13 son esenciales para una buena experimentación. **Todos los experimentos son experimentos diseñados**; desafortunadamente, algunos de ellos se diseñan de manera deficiente y, como resultado, se usan recursos valiosos de manera ineficaz. Los experimentos diseñados estadísticamente permiten eficiencia y economía en el proceso experimental, y la aplicación de métodos estadísticos para examinar los datos resulta en **objetividad científica** al llegar a conclusiones.

12-2 EXPERIMENTO CON UN SOLO FACTOR COMPLETAMENTE ALEATORIZADO

12-2.1 Un ejemplo

Un fabricante de papel para hacer bolsas para comestibles, se encuentra interesado en mejorar la resistencia a la tensión del producto. El departamento de ingeniería del producto piensa que la resistencia a la tensión es una función de la concentración de madera dura en la pulpa y que el rango de las concentraciones de madera dura de interés práctico está entre 5% y 20%. El equipo de ingenieros responsable del estudio decide investigar cuatro niveles de concentración de madera dura: 5%, 10%, 15% y 20%. Deciden hacer seis ejemplares de prueba con cada nivel de

concentración, utilizando una planta piloto. Las 24 muestras se prueban, en orden aleatorio, con una máquina de laboratorio para probar la resistencia. En la tabla 12-1 se muestran los datos de este experimento.

Este es un ejemplo de un experimento con un solo factor completamente aleatorizado con cuatro niveles del factor. A los niveles del factor en ocasiones se les llama **tratamientos**, y cada tratamiento tiene seis observaciones o **réplicas**. El papel de la **aleatorización** en este experimento es de suma importancia. Al aleatorizar el orden de las 24 corridas, se equilibra aproximadamente el efecto de cualquier variable perturbadora que puede influir en la resistencia a la tensión observada. Por ejemplo, suponer que hay un efecto de calentamiento en la máquina para probar la tensión; es decir, entre más tiempo esté funcionando la máquina, mayor será la resistencia a la tensión observada. Si las 24 corridas se hacen en orden ascendente de la concentración de madera dura (es decir, se prueban primero los seis ejemplares con la concentración de 5%, seguidas por los seis ejemplares con la concentración de 10%, etc.), entonces cualquier diferencia observada en la resistencia a la tensión también podría ser causada por el efecto de calentamiento.

Es importante analizar gráficamente los datos de un experimento diseñado. En la figura 12-1*a* se presentan los diagramas de caja de la resistencia a la tensión con los cuatro niveles de concentración de madera dura. Esta figura indica que modificar la concentración de madera dura tiene un efecto sobre la resistencia a la tensión; específicamente, las concentraciones más altas de madera dura producen resistencias a la tensión observadas más altas. Además, la distribución de la resistencia a la tensión en un nivel particular de madera dura es razonablemente simétrica, y la variabilidad de la resistencia a la tensión no sufre cambios radicales cuando la concentración de madera dura cambia.

Siempre es una buena idea hacer la interpretación gráfica de los datos. Los diagramas de caja muestran la variabilidad de las observaciones *dentro* de un tratamiento (nivel del factor) y la variabilidad *entre* los tratamientos. Se comentará ahora cómo pueden analizarse estadísticamente los datos de un experimento aleatorizado con un solo factor.

12-2.2 El análisis de varianza

Suponga que se tienen a diferentes niveles de un factor particular que quieren compararse. En ocasiones, a cada nivel del factor se le llama tratamiento, un término muy general que puede rastrearse hasta las primeras aplicaciones de la metodología del diseño experimental en las ciencias agrícolas. La respuesta para cada uno de los a tratamientos es una variable aleatoria. Los

Tabla 12-1 Resistencia a la tensión del papel (psi)

Concentración	Observaciones							
de madera dura (%)	1	2	3	4	5	6	Totales	Promedios
5	7	8	15	11	9	10	60	10.00
10	12	17	13	18	19	15	94	15.67
15	14	18	19	17	16	18	102	17.00
20	19	25	22	23	18	20	127	21.17
							383	15.96

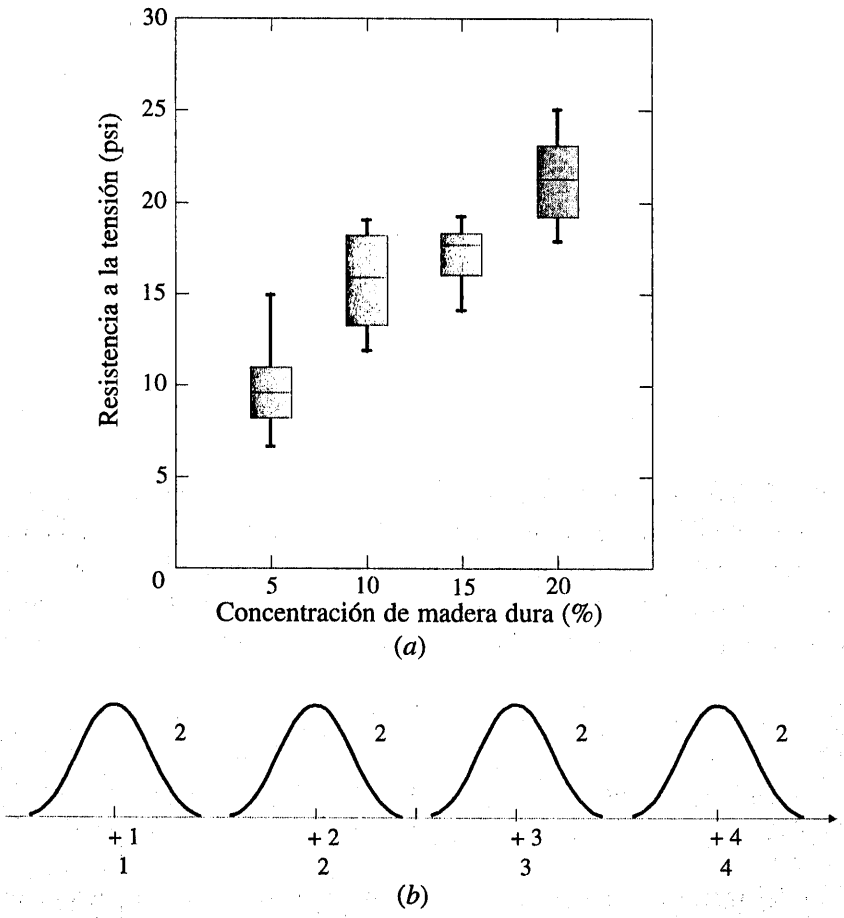


Figura 12-1 ***a***) Diagramas de caja de los datos de la concentración de madera dura. ***b***) Gráfica del modelo de la ecuación 12-1 para el experimento completamente autorizado con un solo factor.

datos observados aparecerían como se muestra en la tabla 12-2. Un dato de la tabla 12-2, por ejemplo y_{ij}, representa la observación j-ésima hecha bajo el tratamiento i-ésimo. Se considera inicialmente el caso en que hay un número igual de observaciones, n, para cada tratamiento.

Las observaciones de la tabla 12-2 pueden describirse con el **modelo estadístico lineal**

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{12-1}$$

donde Y_{ij} es una variable aleatoria que denota la observación ij-ésima, μ es un parámetro común a todos los tratamientos llamado la **media global**, τ_i es un parámetro asociado con el tratamiento i-ésimo llamado el **efecto del tratamiento** i-ésimo, y ε_{ij} es un componente del error aleatorio. Obsérvese que el modelo pudo haberse escrito como

$$Y_{ij} = \mu_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases}$$

Tabla 12-2 Datos típicos de un experimento con un solo factor

Tratamiento	Observaciones				Totales	Promedios
1	y_{11}	y_{12}	...	y_{1n}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
a	y_{a1}	y_{a2}	...	y_{an}	$y_{a\cdot}$	$\bar{y}_{a\cdot}$
					$y_{\cdot\cdot}$	$\bar{y}_{\cdot\cdot}$

donde $\mu_i = \mu + \tau_i$ es la media del tratamiento i-ésimo. En esta forma del modelo, se observa que cada tratamiento define una población que tiene media μ_i, que se compone de la media global μ más un efecto τ_i que se debe a ese tratamiento particular. Se supondrá que los errores εij tienen una distribución normal e independiente con media cero y varianza σ^2. Por lo tanto, cada tratamiento puede considerarse como una población normal con media μ_i y varianza σ^2. Véase la figura 12-1*b*.

La ecuación 12-1 es el modelo fundamental de un experimento con un solo factor. Además, puesto que se requiere que las observaciones se hagan en orden aleatorio y que el medio ambiente (con frecuencia llamado las unidades experimentales) en que se aplican los tratamientos sea tan uniforme como sea posible, a este diseño se le llama **diseño experimental completamente aleatorizado**.

Los niveles a del factor del experimento podrían haberse elegido de dos maneras diferentes. Primera, el experimentador podría haber elegido específicamente los tratamientos a. En esta situación, quieren probarse hipótesis acerca de las medias de los tratamientos, y las conclusiones no pueden generalizarse a tratamientos similares que no se consideraron. Además, quizá quisieran estimarse los **efectos del tratamiento**. A esto se le llama el **modelo de efectos fijos**. De manera alternativa, los tratamientos a podrían ser una muestra aleatoria de una población más grande de tratamientos. En esta situación, sería conveniente estar en posibilidad de extender las conclusiones (las cuales se basan en la muestra de tratamientos) a todos los tratamientos de la población, independientemente de si se consideraron de manera explícita o no en el experimento. En este caso, los efectos del tratamiento τ_i son variables aleatorias, y el conocimiento acerca de los efectos particulares investigados, es relativamente irrelevante. Más bien, se prueban hipótesis acerca de la variabilidad de las τ_i, y se intenta estimar esta variabilidad. A éste se le llama el modelo de **efectos aleatorios**, o **de los componentes de la varianza**.

En esta sección se desarrolla el análisis de varianza para el modelo de efectos fijos. El análisis de varianza no es un tema nuevo; se usó antes en la presentación del análisis de regresión. Sin embargo, en esta sección se indica cómo puede usarse para probar la igualdad de los efectos de los tratamientos. En el modelo de efectos fijos, los efectos del tratamiento τ_i por lo general se definen como desviaciones de la media global μ, por lo que

$$\sum_{i=1}^{a} \tau_i = 0 \tag{12-2}$$

Sea que $y_{i\cdot}$ represente el total de las observaciones bajo el tratamiento i-ésimo y que $\bar{y}_{i\cdot}$ represente el promedio de las observaciones bajo el tratamiento i-ésimo. De manera similar, sea que $y_{\cdot\cdot}$ represente el gran total de todas las observaciones y que $\bar{y}_{\cdot\cdot}$ represente la gran media de todas las observaciones. Expresado matemáticamente

$$y_{i\cdot} = \sum_{j=1}^{n} y_{ij} \quad \bar{y}_{i\cdot} = y_{i\cdot}/n \quad i = 1, 2, \ldots, a$$

$$y_{\cdot\cdot} = \sum_{i=1}^{a}\sum_{j=1}^{a} y_{ij} \quad \bar{y}_{\cdot\cdot} = y_{\cdot\cdot}/N \tag{12-3}$$

donde $N = an$ es el número total de observaciones. Por tanto, la notación del subíndice "punto" implica la operación suma sobre el subíndice que reemplaza.

El interés se encuentra en probar la igualdad de las a medias de los tratamientos $\mu_1, \mu_2, \ldots, \mu_a$. Usando la ecuación 12-2, se encuentra que esto es equivalente a probar las hipótesis

$$\begin{aligned} &H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\ &H_1 : \tau_i \neq 0 \text{ para al menos una } i \end{aligned} \tag{12-4}$$

Por tanto, si la hipótesis nula es verdadera, cada observación consta de la media global μ más un valor que se admite como real del componente del error aleatorio ε_{ij}. Esto es equivalente a decir que la totalidad de las N observaciones se toman de una distribución normal con media μ y varianza σ^2. Por lo tanto, si la hipótesis nula es verdadera, cambiar los niveles del factor no tiene efecto alguno sobre la respuesta media.

En el análisis de varianza se hace la partición de la variabilidad total de los datos muestrales, en dos partes componentes. Así, la prueba de la hipótesis de la ecuación 12-4 se basa en una comparación de dos estimaciones independientes de la varianza poblacional. La variabilidad total de los datos la describe la **suma total de cuadrados**

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{\cdot\cdot})^2$$

La partición de la suma total de cuadrados se da en la siguiente definición.

La **suma de cuadrados identidad** es

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{\cdot\cdot})^2 = n\sum_{i=1}^{a}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\cdot})^2 \tag{12-5}$$

La demostración de esta identidad es directa. Obsérvese que puede escribirse

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2=\sum_{i=1}^{a}\sum_{j=1}^{n}[(\bar{y}_{i.}-\bar{y}_{..})+(y_{ij}-\bar{y}_{i.})]^2$$

o bien

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2=n\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})^2+\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2$$
$$+2\sum_{i=1}^{a}\sum_{j=1}^{n}(\bar{y}_{i.}-\bar{y}_{..})(y_{ij}-\bar{y}_{i.}) \qquad (12\text{-}6)$$

Obsérvese que el término del producto cruzado de la ecuación 12-6 es cero, ya que

$$\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})=y_{i.}-n\bar{y}_{i.}=y_{i.}-n(y_{i.}/n)=0$$

Por lo tanto, se ha demostrado que la ecuación 12-6 se reduce a la ecuación 12-5.

La identidad de la ecuación 12-5 muestra que puede hacerse la partición de la variabilidad total de los datos, medida por la suma total de cuadrados corregida, en una suma de cuadrados de las diferencias entre las medias de los tratamientos y la gran media, y una suma de cuadrados de las diferencias de las observaciones dentro de un tratamiento con la media del tratamiento. Las diferencias entre las medias observadas de los tratamientos y la gran media, miden las diferencias entre los tratamientos, mientras que las diferencias de las observaciones dentro de un tratamiento, de la media del tratamiento, pueden deberse tan sólo al error aleatorio. Por lo tanto, usando símbolos la ecuación 12-5 puede escribirse como

$$SS_T = SS_{\text{Tratamientos}} + SS_E \qquad (12\text{-}7)$$

donde

$$SS_T=\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2 = \text{ suma total de cuadros}$$

$$SS_{\text{Tratamientos}} = n\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})^2 = \text{suma de cuadrados de los tratamientos}$$

y

$$SS_E=\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2 = \text{suma de cuadros del error}$$

Es posible profundizar en la forma en que el análisis de varianza opera examinando los valores esperados de $SS_{\text{Tratamientos}}$ y SS_E. Esto llevará a un estadístico apropiado para probar la hipótesis de que no hay diferencias entre las medias de los tratamientos (o bien $\tau_i = 0$).

El valor esperado de la suma de cuadrados de los tratamientos es

$$E(SS_{\text{Tratamientos}}) = (a-1)\sigma^2 + n\sum_{i=1}^{a}\tau_i^2$$

La demostración de esta expresión es sencilla. Al aplicar directamente el operador del valor esperado a $SS_{\text{Tratamientos}}$ se obtiene

$$E(SS_{\text{Tratamientos}}) = E\left[n\sum_{i=1}^{a}(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2\right]$$

y por el modelo $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ se tiene

$$\bar{Y}_{i\cdot} = \mu + \tau_i + \bar{\varepsilon}_{i\cdot}$$

y

$$\bar{Y}_{\cdot\cdot} = \mu + \bar{\varepsilon}$$

puesto que $\sum_{i=1}^{a}\tau_i = 0$. Al sustituir $\bar{Y}_{i\cdot}$ y $\bar{Y}_{\cdot\cdot}$ en la expresión para $SS_{\text{Tratamientos}}$ se obtiene

$$E(SS_{\text{Tratamientos}}) = E\left[n\sum_{i=1}^{a}(\tau_i + \bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})^2\right]$$

$$= E\left[n\sum_{i=1}^{a}\tau_i^2 + n\sum_{i=1}^{a}\bar{\varepsilon}_{i\cdot}^2 + an\bar{\varepsilon}_{\cdot\cdot}^2 + 2n\sum_{i=1}^{a}\tau_i\bar{\varepsilon}_{i\cdot} - 2n\bar{\varepsilon}_{\cdot\cdot}\sum_{i=1}^{a}\tau_i - 2n\bar{\varepsilon}_{\cdot\cdot}\sum_{i=1}^{a}\bar{\varepsilon}_{i\cdot}\right]$$

Sin embargo, puesto que las ε_{ij} son variables aleatorias independientes con media cero y varianza σ^2, se encuentra que

$$E(\bar{\varepsilon}_{i\cdot}^2) = \frac{\sigma^2}{n}, \quad E(\bar{\varepsilon}_{\cdot\cdot}^2) = \frac{\sigma^2}{an} \text{ y } E(\bar{\varepsilon}_{i\cdot}) = 0$$

Por lo tanto

$$E(SS_{\text{Tratamientos}}) = n\sum_{i=1}^{a}\tau_i^2 + a\sigma^2 - \sigma^2$$

$$= (a-1)\sigma^2 + n\sum_{i=1}^{a}\tau_i^2$$

Ahora bien, si la hipótesis nula de la ecuación 12-4 es verdadera, cada τ_i es igual a cero y

$$E\left(\frac{SS_{\text{Tratamientos}}}{a-1}\right)=\sigma^2$$

Si la hipótesis alternativa es verdadera, entonces

$$E\left(\frac{SS_{\text{Tratamientos}}}{a-1}\right)=\sigma^2+\frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

Al cociente $MS_{\text{Tratamientos}} = SS_{\text{Tratamientos}}/(a-1)$ se le llama el **cuadrado medio de los tratamientos**. Por tanto, si H_0 es verdadera, $MS_{\text{Tratamientos}}$ es un estimador insesgado de σ^2, mientras que si H_1 es verdadera, $MS_{\text{Tratamientos}}$ es una estimación de σ^2 más un término positivo que incorpora la variación debida a la diferencia sistemática en las medias de los tratamientos.

Utilizando un enfoque similar, puede demostrarse que el valor esperado de la suma de cuadrados del error es $E(SS_E) = a(n-1)\sigma^2$. Por lo tanto, el **cuadrado medio de error** $MS_E = SS_E/[a(n-1)]$ es un estimador insesgado de σ^2 independientemente de si H_0 es verdadera o no.

El **cuadrado medio de error**

$$MS_E = \frac{SS_E}{a(n-1)}$$

es un estimador insesgado de σ^2.

También hay una partición del número de grados de libertad que corresponde a la suma de cuadrados identidad de la ecuación 12-5. Es decir, hay $an = N$ observaciones; por tanto, SS_T tiene $an-1$ grados de libertad. Como hay a niveles del factor, $SS_{\text{Tratamientos}}$ tiene $a-1$ grados de libertad. Por último, dentro de cualquier tratamiento hay n réplicas que proporcionan $n-1$ grados de libertad con los cuales estimar el error experimental. Puesto que hay a tratamientos, se tienen $a(n-1)$ grados de libertad para el error. Por lo tanto, la partición de los grados de libertad es

$$an-1 = a-1+a(n-1)$$

Suponga ahora que cada una de las a poblaciones puede modelarse como una distribución normal. Usando este supuesto puede demostrarse que si la hipótesis nula H_0 es verdadera, el cociente

$$F_0=\frac{SS_{\text{Tratamientos}}/(a-1)}{SS_E/[a(n-1)]}=\frac{MS_{\text{Tratamientos}}}{MS_E} \tag{12-8}$$

tiene una distribución F con $a - 1$ y $a(n - 1)$ grados de libertad. Además, por los cuadrados medios esperados, se sabe que MS_E es un estimador insesgado de σ^2. Asimismo, bajo la hipótesis nula, $MS_{\text{Tratamientos}}$ es un estimador insesgado de σ^2. Sin embargo, si la hipótesis nula es falsa, entonces el valor esperado de $MS_{\text{Tratamientos}}$ es mayor que σ^2. Por lo tanto, bajo la hipótesis alternativa, el valor esperado del numerador del estadístico de la prueba (ecuación 12-8) es mayor que el valor esperado del denominador. Por consiguiente, H_0 deberá rechazarse si el estadístico es grande. Esto implica una región crítica de una cola y de cola superior. Por lo tanto, H_0 se rechazaría si $f_0 > f_{\alpha,\, a-1,\, a(n-1)}$ donde f_0 se calcula con la ecuación 12-8.

Es posible obtener fórmulas de cálculo eficientes para las sumas de los cuadrados desarrollando y simplificando las definiciones de $SS_{\text{Tratamientos}}$ y SS_T. Se obtienen así los resultados siguientes.

Definición

Las fórmulas para calcular las sumas de cuadrados del análisis de varianza con tamaños de las muestras iguales en cada tratamiento son

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{12-9}$$

y

$$SS_{\text{Tratamientos}} = \sum_{i=1}^{a} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N} \tag{12-10}$$

La suma de cuadrados del error se obtiene por sustracción como

$$SS_E = SS_T - SS_{\text{Tratamientos}} \tag{12-11}$$

Los cálculos para este procedimiento de prueba por lo general se resumen en forma tabular como se muestra en la tabla 12-3. A ésta se le llama **tabla de análisis de varianza**.

Tabla 12-3 El análisis de varianza para un experimento con un solo factor, modelo de efectos físicos

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$SS_{\text{Tratamientos}}$	$a - 1$	$MS_{\text{Tratamientos}}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
Error	SS_E	$a(n - 1)$	MS_E	
Total	SS_T	$an - 1$		

EJEMPLO 12-1

Considérese el experimento de la resistencia a la tensión del papel descrito en la sección 12-2.1. Puede usarse el análisis de varianza para probar la hipótesis de que diferentes concentraciones de madera dura no afectan la resistencia a la tensión media del papel.

Las hipótesis son

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$$

$$H_1 : \tau_i \neq 0 \text{ para al menos una } i$$

Se usará $\alpha = 0.01$. Las sumas de cuadrados para el análisis de varianza se calculan con las ecuaciones 12-9, 12-10 y 12-11 de la siguiente manera:

$$\begin{aligned} SS_T &= \sum_{i=1}^{4}\sum_{j=1}^{6} y_{ij}^2 - \frac{y_{..}^2}{N} \\ &= (7)^2 + (8)^2 + \cdots + (20)^2 - \frac{(383)^2}{24} = 512.96 \\ SS_{\text{Tratamientos}} &= \sum_{i=1}^{4} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N} \\ &= \frac{(60)^2 + (94)^2 + (102)^2 + (127)^2}{6} - \frac{(383)^2}{24} = 382.79 \\ SS_E &= SS_T - SS_{\text{Tratamientos}} \\ &= 512.96 - 382.79 = 130.17 \end{aligned}$$

En la tabla 12-4 se resume el análisis de varianza. Puesto que $f_{0.01,3,20} = 4.94$, se rechaza H_0 y se concluye que la concentración de madera dura de la pulpa afecta de manera significativa la resistencia del papel. También puede encontrarse un valor P para este estadístico de la prueba como sigue:

$$P = P(F_{3,20} > 19.60) \simeq 3.59 \times 10^{-6}$$

Puesto que el valor $P \simeq 3.59 \times 10^{-6}$ es considerablemente más pequeño que $\alpha = 0.01$, se cuenta con evidencia sólida para concluir que H_0 no es verdadera.

Tabla 12-4 Análisis de varianza para los datos de la resistencia a la tensión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Concentración de madera dura	382.79	3	127.60	19.60	3.59 E-6
Error	130.17	20	6.51		
Total	512.96	23			

En algunos experimentos con un solo factor, el número de observaciones que se hacen bajo cada tratamiento puede ser diferente. Se dice entonces que el diseño es **desbalanceado**. El análisis de varianza descrito arriba sigue siendo válido, aunque es necesario hacer ligeras modificaciones en las fórmulas de las sumas de cuadrados. Sea que se hagan n_i observaciones bajo el tratamiento i ($i = 1, 2, ..., a$), y sea $N = \sum_{i=1}^{a} n_i$. el número total de observaciones. Las fórmulas para calcular SS_T y $SS_{\text{Tratamientos}}$ se presentan en la siguiente definición.

Definición

Las fórmulas para calcular las sumas de cuadrados para el análisis de varianza con tamaños de las muestras n_i diferentes en cada tratamiento son

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SS_{\text{Tratamientos}} = \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N}$$

y

$$SS_E = SS_T - SS_{\text{Tratamientos}}$$

Elegir un diseño balanceado tiene dos ventajas importantes. Primera, el procedimiento de prueba es relativamente insensible a las desviaciones pequeñas del supuesto de la igualdad de las varianzas si los tamaños de las muestras son iguales. Este no es el caso para tamaños de las muestras diferentes. Segunda, la potencia de la prueba se maximiza si las muestras son de tamaños iguales.

12-2.3 Salida de computadora

Muchos paquetes de software tienen la capacidad de analizar datos de experimentos diseñados utilizando el análisis de varianza. En la tabla 12-5 se presenta la salida de la rutina del análisis de varianza simple del programa Minitab para el experimento de la resistencia a la tensión descrito en el ejemplo 12-1. Los resultados concuerdan bastante bien con los cálculos manuales presentados previamente en la tabla 12-4.

La salida de Minitab también incluye intervalos de confianza de 95% para la media de cada tratamiento individual. La media del tratamiento i-ésimo se define como

$$\mu_i = \mu + \tau_i \qquad i = 1, 2, ..., a$$

Un estimador puntual de μ_i es $\hat{\mu}_i = \bar{Y}_{i.}$. Entonces, si se supone que los errores tienen una distribución normal, el promedio de cada tratamiento tiene una distribución normal con media μ_i y

Tabla 12-5 Salida del programa Minitab para el análisis de varianza del ejemplo 12-1

Análisis de varianza de un solo factor

Análisis de varianza de la resistencia

Fuente	DF	SS	MS	F	P
Concentración	3	382.79	127.60	19.61	0.000
Error	20	130.17	6.51		
Total	23	512.96			

Nivel	N	Media	Desviación estándar
5	6	10.000	2.828
10	6	15.667	2.805
15	6	17.000	1.789
20	6	21.167	2.639

Intervalos de confianza individuales de 95% para la media basados en la desviación estándar agrupada

```
-----+---------+---------+---------+-
(---*---)
          (---*---)
             (---*---)
                     (---*---)
-----+---------+---------+---------+-
    10.0      15.0      20.0      25.0
```

Desviación estándar agrupada = 2.551

Comparaciones por pares de Fisher

Índice de error de familia = 0.192
Índice de error individual = 0.0500

Valor crítico = 2.086

Intervalos para (media nivel columna) – (media nivel renglón)

	5	10	15
10	–8.739 –2.594		
15	–10.072 –3.928	–4.406 1.739	
20	–14.239 –8.904	–8.572 –2.428	–7.239 –1.094

varianza σ^2/n. Por tanto, si se conociera σ^2, podría usarse la distribución normal para construir un intervalo de confianza. Usando MS_E como estimador de σ^2 (esta es la "Desviación estándar agrupada" que aparece en la salida de Minitab) el intervalo de confianza se basaría en la distribución t, ya que

$$T = \frac{\bar{Y}_{i\cdot} - \mu_i}{\sqrt{MS_E / n}}$$

tiene una distribución t con $a(n-1)$ grados de libertad. Esto lleva a la siguiente definición del intervalo de confianza.

Definición

Un intervalo de confianza del $100(1-\alpha)$ por ciento para la media del tratamiento i-ésimo, μ_i, es

$$\bar{y}_{i\cdot} - t_{\alpha/2,\, a(n-1)}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_{i\cdot} + t_{\alpha/2,\, a(n-1)}\sqrt{\frac{MS_E}{n}} \qquad (12\text{-}12)$$

La ecuación 12-12 se usa para calcular los intervalos de confianza indicados gráficamente en la salida del programa Minitab de la tabla 12-5.

EJEMPLO 12-2

Las estimaciones de la media de la resistencia a la tensión para el experimento del ejemplo 12-1 son las siguientes:

5% concentración de madera dura: $\hat{\mu}_1 = \bar{y}_{1\cdot} =$ 10.00 psi

10% concentración de madera dura: $\hat{\mu}_2 = \bar{y}_{2\cdot} =$ 15.67 psi

15% concentración de madera dura: $\hat{\mu}_3 = \bar{y}_{3\cdot} =$ 17.00 psi

20% concentración de madera dura: $\hat{\mu}_4 = \bar{y}_{4\cdot} =$ 21.17 psi

El intervalo de confianza de 95% para la resistencia a la tensión media con 20% de madera dura se encuentra con la ecuación 12-12 como sigue:

$$\left[\bar{y}_{4\cdot} \pm t_{0.025,\,20}\sqrt{MS_E/n}\right]$$
$$\left[21.17 \pm (2.086)\sqrt{6.51/6}\right]$$
$$[21.17 \pm 2.17]$$

Por lo tanto, el intervalo de confianza deseado es

$$19.00 \text{ psi} \le \mu_4 \le 23.34 \text{ psi}$$

Los otros intervalos de confianza para la concentración de madera dura de 5%, 10% y 15% se calculan de manera similar. El examen visual de los intervalos de confianza de la salida de Minitab sugieren que las resistencias a la tensión medias con 10% y 15% de madera dura son similares.

También es de interés encontrar intervalos de confianza para la diferencia de las medias de dos tratamientos, por ejemplo, $\mu_i - \mu_j$. El estimador puntual de $\mu_i - \mu_j$ es $\bar{Y}_{i\cdot} - \bar{Y}_{j\cdot}$, y la varianza de este estimador es

$$V(\bar{Y}_{i\cdot} - \bar{Y}_{j\cdot}) = \frac{\sigma^2}{n} + \frac{\sigma^2}{n} = \frac{2\sigma^2}{n}$$

Ahora bien, si se usa MS_E para estimar σ^2, entonces

$$T = \frac{Y_{i\cdot} - Y_{j\cdot} - (\mu_i - \mu_j)}{\sqrt{2MS_E/n}}$$

tiene una distribución t con $a(n-1)$ grados de libertad. Por lo tanto, un intervalo de confianza para $\mu_i - \mu_j$ puede basarse en la distribución t.

Definición

Un intervalo de confianza del $100(1-\alpha)$ por ciento para la diferencia de las medias de dos tratamientos $\mu_i - \mu_j$ es

$$\bar{y}_{i\cdot} - \bar{y}_{j\cdot} - t_{\alpha/2,\,a(n-1)}\sqrt{\frac{2MS_E}{n}} \le \mu_i - \mu_j \le \bar{y}_{i\cdot} - \bar{y}_{j\cdot} + t_{\alpha/2,\,a(n-1)}\sqrt{\frac{2MS_E}{n}} \qquad (12\text{-}13)$$

Un intervalo de confianza de 95% para la diferencia de las medias $\mu_3 - \mu_2$ se calcula con la ecuación 12-13 como sigue:

$$\left[\bar{y}_{3\cdot} - \bar{y}_{2\cdot} \pm t_{0.025,\,20}\sqrt{2MS_E/n}\right]$$
$$\left[17.00 - 15.67 \pm (2.086)\sqrt{2(6.51)/6}\right]$$
$$[17.00 - 15.67 \pm 3.07]$$

Por tanto, el intervalo de confianza de 95% para $\mu_3 - \mu_2$ es

$$-1.74 \le \mu_3 - \mu_2 \le 4.40$$

Puesto que el intervalo de confianza incluye al cero, se concluiría que no hay diferencia en la resistencia a la tensión media en estos dos niveles de madera dura particulares.

La parte inferior de la salida de computadora de la tabla 12-5 proporciona información adicional respecto a cuáles medias específicas son diferentes. Se revisará esto con mayor detalle en la sección 12-3.

12-2.4 Análisis de residuales y verificación del modelo

En el análisis de varianza del modelo simple o de un solo factor, se supone que las observaciones siguen una distribución normal e independiente con la misma varianza para cada tratamiento o nivel del factor. Estos supuestos deberán verificarse examinando los residuales. Un residual es la diferencia entre una observación y_{ij} y su valor estimado (o ajustado) en el modelo estadís-

Tabla 12-6 Residuales para el experimento de la resistencia a la tensión

Concentración de madera dura	Residuales					
5%	−3.00	−2.00	5.00	1.00	−1.00	0.00
10%	−3.67	1.33	−2.67	2.33	3.33	−0.67
15%	−3.00	1.00	2.00	0.00	−1.00	1.00
20%	−2.17	3.83	0.83	1.83	−3.17	−1.17

tico bajo estudio, denotado como $\hat{y}_{ij}$. Para el diseño completamente aleatorizado $\hat{y}_{ij} = \bar{y}_{i\cdot}$ y cada residual es $e_{ij} = y_{ij} - \bar{y}_{i\cdot}$, es decir, la diferencia entre una observación y la media observada del tratamiento correspondiente. En la tabla 12-6 se muestran los residuales para el experimento del porcentaje de madera dura. Al utilizar $\bar{y}_{i\cdot}$ para calcular cada residual en esencia, se elimina el efecto de la concentración de madera dura de esos datos; por consiguiente, los residuales contienen información acerca de la variabilidad no explicada.

El supuesto de normalidad puede verificarse construyendo una gráfica de probabilidad normal de los residuales. Para verificar el supuesto de la igualdad de las varianzas en cada nivel del factor, se grafican los residuales contra los niveles del factor y se compara la dispersión de los residuales. También es conveniente graficar los residuales contra $\bar{y}_{i\cdot}$ (llamada en ocasiones el valor ajustado); la variabilidad de los residuales no deberá depender en forma alguna del valor de $\bar{y}_{i\cdot}$. La mayoría de los paquetes de software de estadística construirán estas gráficas si así se estipula. Cuando aparece un patrón en estas gráficas, por lo general sugiere la necesidad de una transformación, es decir, de analizar los datos con una métrica diferente. Por ejemplo, si la variabilidad de los residuales se incrementa con $\bar{y}_{i\cdot}$, entonces debería considerarse una transformación como $\log y$ o $\sqrt{y}$. En algunos problemas, la dependencia de la dispersión de los residuales con la $\bar{y}_{i\cdot}$ media observada, constituye información de suma utilidad. Puede ser deseable seleccionar el nivel del factor que dé como resultado la respuesta máxima; sin embargo, este nivel también puede producir una variación mayor en la respuesta de una corrida a otra.

El supuesto de independencia puede verificarse graficando los residuales contra el tiempo o contra el orden de las corridas en que se realizó el experimento. Un patrón en esta gráfica, como secuencias de residuales positivos y negativos, puede indicar que las observaciones no son independientes. Esto sugiere que el tiempo o el orden de las corridas es importante o que las variables que cambian con el tiempo son importantes y no se han incluido en el diseño experimental.

En la figura 12-2 se muestra la gráfica de probabilidad normal de los residuales del experimento de la resistencia a la tensión del papel. En las figuras 12-3 y 12-4 se presentan los residuales graficados contra los niveles del factor y el valor ajustado $\bar{y}_{i\cdot}$, respectivamente. Estas gráficas no revelan ninguna falta de adecuación del modelo ni algún problema excepcional con los supuestos.

EJERCICIOS DE LA SECCIÓN 12-2

12-1. En *Design and Analysis of Experiments*, 4a. edición (John Wiley & Sons), D.C. Montgomery describe un experimento en el que la resistencia a la tensión de una fibra sintética es de interés para el fabricante. Se piensa que la resistencia se relaciona con el porcentaje

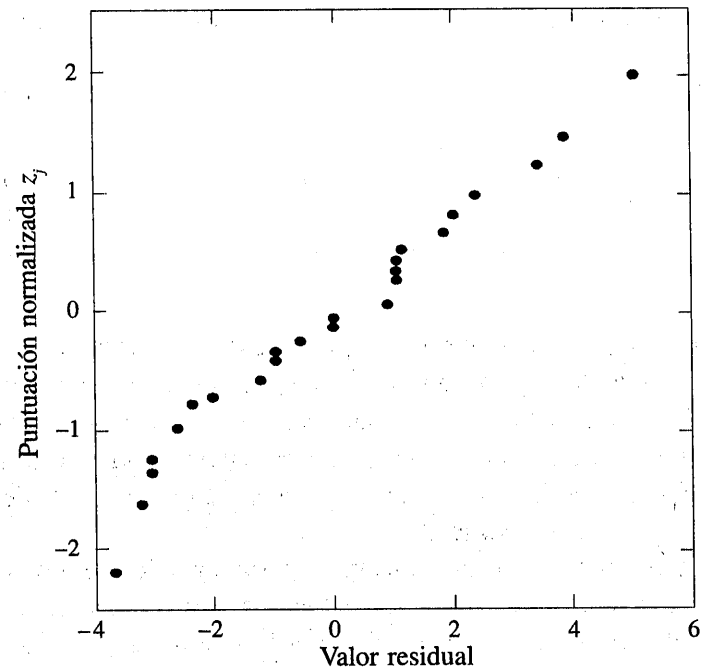


Figura 12-2 Gráfica de probabilidad normal de los residuales del experimento de la concentración de madera dura.

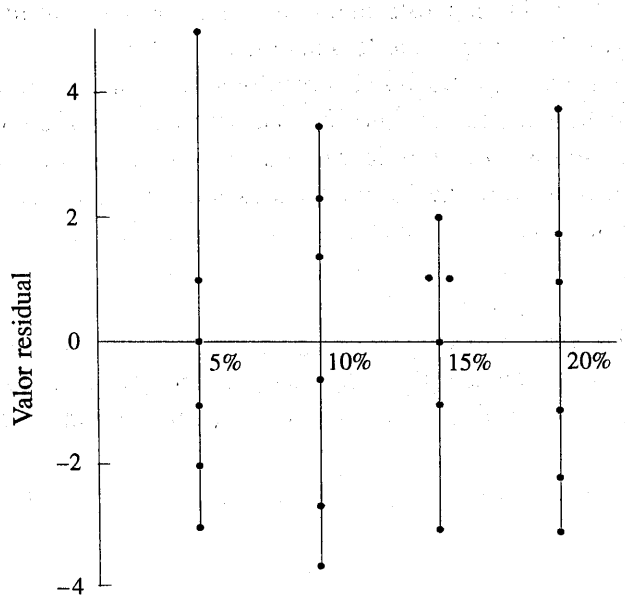


Figura 12-3 Gráfica de los residuales contra los niveles del factor (concentración de madera dura).

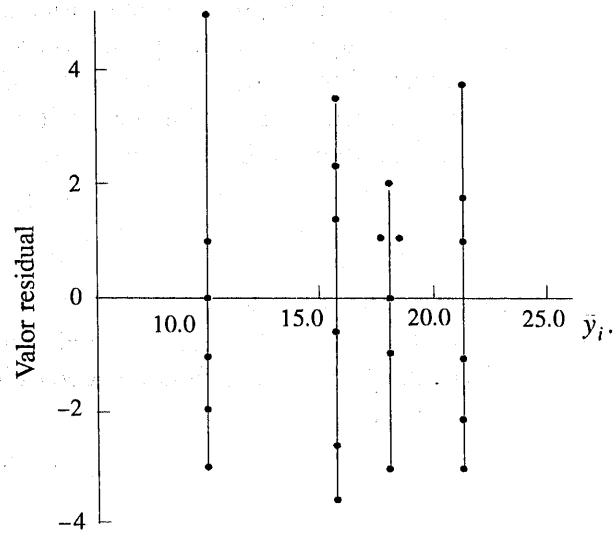


Figura 12-4 Gráfica de los residuales contra $\bar{y}_{i\cdot}$.

de algodón de la fibra. Se usan cinco niveles del porcentaje de algodón, y se hacen cinco réplicas en orden aleatorio, obteniéndose los siguientes datos como resultado:

Porcentaje de algodón	Observaciones 1	2	3	4	5
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11

a) ¿El porcentaje de algodón afecta la resistencia a la ruptura? Trace diagramas de caja comparativos y haga un análisis de varianza. Use $\alpha = 0.05$.

b) Grafique la resistencia a la tensión promedio contra el porcentaje de algodón e interprete los resultados.

c) Encuentre los residuales y examínelos en cuanto a la adecuación del modelo.

12-2. En "Diseño ortogonal para optimización de procesos y su aplicación en el grabado químico con plasma" (*Solid State Technology*), G.Z. Yin y D.W. Jillie describen un experimento para determinar el efecto de la velocidad de flujo del C_2F_6 sobre la uniformidad del grabado químico en una oblea de silicio usada para fabricar circuitos integrados. Se usan tres velocidades de flujo en el experimento, y la uniformidad resultante (en por ciento) de seis réplicas se muestra abajo.

Flujo de C_2F_6	Observaciones 1	2	3	4	5	6
125	2.7	4.6	2.6	3.0	3.2	3.8
160	4.9	4.6	5.0	4.2	3.6	4.2
200	4.6	3.4	2.9	3.5	4.1	5.1

a) ¿La velocidad de flujo del C_2F_6 afecta la uniformidad del grabado químico? Construya diagramas de caja para comparar los niveles del factor y haga un análisis de varianza. Use $\alpha = 0.05$.

b) ¿Los residuales indican algún problema con los supuestos fundamentales?

12-3. Se estudia la resistencia a la compresión del concreto, y se investigan cuatro técnicas de mezclado diferentes. Se han obtenido los siguientes datos:

Técnica de mezclado	Resistencia a la compresión (psi)			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

a) Pruebe la hipótesis de que las técnicas de mezclado afectan la resistencia a la compresión del concreto. Use $\alpha = 0.05$.

b) Encuentre el valor P para el estadístico F calculado en el inciso a).

c) Analice los residuales de este experimento.

12-4. Se realizó un experimento para determinar si cuatro temperaturas de cocción específicas afectan la densidad de cierto tipo de tabique. El experimento llevó a los siguientes datos:

Temperatura (°F)	Densidad						
100	21.8	21.9	21.7	21.6	21.7	21.5	21.8
125	21.7	21.4	21.5	21.5	—	—	—
150	21.9	21.8	21.8	21.6	21.5	—	—
175	21.9	21.7	21.8	21.7	21.6	21.8	—

a) ¿La temperatura de cocción afecta la densidad de los tabiques? Use $\alpha = 0.05$.

b) Encuentre el valor P para el estadístico F calculado en el inciso a).

c) Analice los residuales de este experimento.

12-5. Un ingeniero electrónico está interesado en el efecto que tienen cinco tipos diferentes de recubrimiento de un cinescopio, sobre la conductividad del mismo en un dispositivo de visualización de un sistema de telecomunicaciones. Se obtuvieron los siguientes datos para la conductividad:

Tipo de recubrimiento	Conductividad			
1	143	141	150	146
2	152	149	137	143
3	134	133	132	127
4	129	127	132	129
5	147	148	144	142

a) ¿Hay alguna diferencia en la conductividad causada por el tipo de recubrimiento? Use $\alpha = 0.01$.
b) Analice los residuales de este experimento.
c) Construya una estimación de un intervalo de 95% para la media del recubrimiento tipo 1. Construya una estimación de un intervalo de 99% para la diferencia entre las medias de los recubrimientos tipo 1 y 4.

12-6. Se determinó el tiempo de respuesta en milisegundos de tres diferentes tipos de circuitos de una calculadora electrónica. Los resultados fueron los siguientes:

Tipo de circuito	Respuesta				
1	19	22	20	18	25
2	20	21	33	27	40
3	16	15	18	26	17

a) Usando $\alpha = 0.01$, pruebe la hipótesis de que los tres tipos de circuitos tienen el mismo tiempo de respuesta.
b) Analice los residuales de este experimento.
c) Encuentre un intervalo de confianza de 95% para el tiempo de respuesta del circuito tres.

12-7. En un artículo de *ACI* (Vol. 84, pp.213-216) *Materials Journal* se describen varios experimentos para investigar el varillado del concreto con una varilla, a fin de liberar el aire atrapado. Se utilizó un cilindro de 3 × 6 pulgadas, y el número de veces que se usó esta varilla es la variable de diseño. La resistencia a la compresión resultante del ejemplar de prueba de concreto es la variable de respuesta. Los datos se muestran en la siguiente tabla.

Nivel de varillado	Resistencia a la compresión		
10	1530	1530	1440
15	1610	1650	1500
20	1560	1730	1530
25	1500	1490	1510

a) ¿Hay alguna diferencia en la resistencia a la compresión debida al nivel de varillado?
b) Encuentre el valor P para el estadístico F del inciso a).
c) Analice los residuales de este experimento. ¿Qué conclusiones pueden sacarse acerca de los supuestos fundamentales del modelo?

12-8. En un artículo de *Environment International* (Vol. 18, No. 4) se describe un experimento en el que se investigó la cantidad de radón liberado en las duchas. Se usó agua enriquecida con radón, y se probaron seis diámetros diferentes de los orificios de las regaderas. Los datos del experimento se presentan en la siguiente tabla.

Diámetro de los orificios	Radón liberado (%)			
0.37	80	83	83	85
0.51	75	75	79	79
0.71	74	73	76	77
1.02	67	72	74	74
1.40	62	62	67	69
1.99	60	61	64	66

a) ¿El tamaño de los orificios afecta el porcentaje promedio de radón liberado? Use $\alpha = 0.05$.
b) Encuentre el valor P para el estadístico F del inciso a).
c) Analice los residuales de este experimento.
d) Encuentre un intervalo de confianza de 95% para el porcentaje promedio de radón liberado cuando el diámetro de los orificios es 1.40.

12-9. En un artículo de *Journal of the Association of Asphalt Paving Technologists* (Vol. 59) se

describe un experimento para determinar el efecto de los vacíos de aire sobre la resistencia porcentual conservada del asfalto. Para los fines del experimento, los vacíos de aire se controlan en tres niveles: bajo (2-4%), medio (4-6%) y alto (6-8%). Los datos se presentan en la tabla siguiente:

Nivel del vacío de aire	Resistencia conservada (%)							
Bajo	106	90	103	90	79	88	92	95
Medio	80	69	94	91	70	83	87	83
Alto	78	80	62	69	76	85	69	85

a) ¿Los diferentes niveles de los vacíos de aire afectan de manera significativa la resistencia conservada promedio? Use $\alpha = 0.01$.
b) Encuentre el valor P para el estadístico F del inciso a).
c) Analice los residuales de este experimento.
d) Encuentre un intervalo de confianza de 95% para la resistencia conservada promedio cuando hay un nivel alto de vacíos de aire.
e) Encuentre un intervalo de confianza de 95% para la diferencia de la resistencia conservada promedio en los niveles bajo y alto de vacíos de aire.

12-10. Un artículo de *Materials Research Bulletin* (Vol. 26, No. 11) investigó cuatro métodos diferentes para preparar el compuesto superconductor $PbMo_6S_8$. Los autores sostienen que la presencia de oxígeno durante el proceso de preparación afecta la temperatura de transición de superconducción T_c del material. Los métodos de preparación 1 y 2 usan técnicas que están diseñadas para eliminar la presencia de oxígeno, mientras que los métodos 3 y 4 permiten la presencia de oxígeno. Se hicieron cinco observaciones de T_c (en °K) para cada método, y los resultados son los siguientes:

Método de preparación	Temperatura de transición T_c(°K)				
1	14.8	14.8	14.7	14.8	14.9
2	14.6	15.0	14.9	14.8	14.7
3	12.7	11.6	12.4	12.7	12.1
4	14.2	14.4	14.4	12.2	11.7

a) ¿Hay evidencia que apoye la afirmación de que la presencia de oxígeno durante la preparación afecta la temperatura de transición media? Use $\alpha = 0.05$.
b) ¿Cuál es el valor P para la prueba F del inciso a)?
c) Analice los residuales de este experimento.
d) Encuentre un intervalo de confianza de 95% para T_c media cuando se usa el método 1 para preparar el material.

12-3 PRUEBAS SOBRE MEDIAS DE TRATAMIENTOS INDIVIDUALES

Supóngase que el análisis de varianza indica que la hipótesis nula deberá rechazarse. Esto implica, desde luego, diferencias entre las medias de los tratamientos, pero sin especificar exactamente cuáles medias son diferentes. Evidentemente, en una situación así se justifica la comparación adicional de las medias de los tratamientos. A los procedimientos para comparar las medias de tratamientos individuales en el análisis de varianza por lo general se les llama métodos de **comparación múltiple**. Estos métodos se introducen en esta sección.

12-3.1 Comparación gráfica de medias

Es sencillo hacer la comparación gráfica de las medias de los tratamientos, después de haber realizado el análisis de varianza. Suponga que el factor tiene a niveles y que $\bar{y}_{1\cdot}$, $\bar{y}_{2\cdot}$, ..., $\bar{y}_{a\cdot}$ son

los promedios observados de estos niveles del factor. Todos los promedios de los tratamientos tienen desviación estándar $\sigma/\sqrt{n}$, donde σ es la desviación estándar de una observación individual. Si todos los promedios de los tratamientos son iguales, las medias observadas $\bar{y}_{i\cdot}$ se comportarían como si fueran un conjunto de observaciones tomadas al azar de una distribución normal con media μ y desviación estándar $\sigma/\sqrt{n}$.

Imaginar que esta distribución normal puede deslizarse sobre un eje abajo del cual se grafican las medias de los tratamientos $\bar{y}_{1\cdot}$, $\bar{y}_{2\cdot}$, ..., $\bar{y}_{a\cdot}$. Si todas las medias de los tratamientos son iguales, deberá haber alguna posición para esta distribución que haga evidente que los valores $\bar{y}_{i\cdot}$ se tomaron de la misma distribución. Si no es este el caso, entonces los valores $\bar{y}_{i\cdot}$ que no parecen haberse sacado de esta distribución, se asocian con tratamientos que producen respuestas medias diferentes.

La única falla en este razonamiento es que σ no se conoce. Sin embargo, puede usarse $\sqrt{MS_E}$ del análisis de varianza para estimar σ. Esto implica que deberá usarse una distribución t en lugar de la normal al hacer la gráfica, pero dada la gran semejanza entre la distribución t y la normal, trazar una curva normal que tenga aproximadamente $6\sqrt{MS_E/n}$ unidades de ancho funcionará bastante bien en la mayoría de los casos.

En la figura 12-5 se ilustra esta situación para el experimento de la concentración de madera dura de la sección 12-2.1. La desviación estándar de esta distribución normal es

$$\sqrt{MS_E/n} = \sqrt{6.51/6} = 1.04$$

Si se visualiza el deslizamiento de esta distribución sobre el eje horizontal, se observa que no hay ninguna posición de esta distribución que sugiera que las cuatro observaciones (las medias graficadas) son valores típicos seleccionados al azar de esa distribución. Esto, desde luego, sería de esperarse, ya que el análisis de varianza ha indicado que las medias difieren, y la ilustración de la figura 12-5 es tan sólo una representación gráfica de los resultados del análisis de varianza. La figura indica que el tratamiento 4 (20% de madera dura) produce el papel con la resistencia a la tensión media más alta que los otros tratamientos, y que el tratamiento 1 (5% de madera dura) produce la resistencia a la tensión media más baja que los tratamientos restantes. Las medias de los tratamientos 2 y 3 (10 y 15% de madera dura, respectivamente) no difieren.

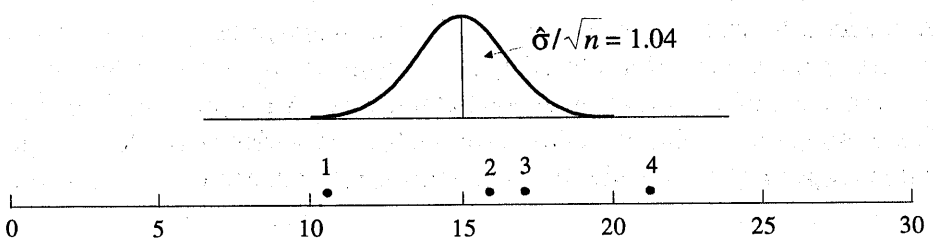


Figura 12-5 Promedio de la resistencia a la tensión del experimento de la concentración de madera dura en relación con una distribución normal de desviación estándar $\sqrt{MS_E/n} = \sqrt{6.51/6} = 1.04$.

Este procedimiento es una técnica de comparación múltiple aproximada pero muy eficaz. Sin embargo, hay procedimientos de prueba con fundamentos estadísticos más formales. Para mayores detalles sobre estos procedimientos, véase Montgomery. Se describen a continuación de manera sucinta dos de estos procedimientos: el método de contrastes ortogonales y el de la diferencia significativa mínima de Fisher.

12-3.2 Contrastes ortogonales

Muchos de los procedimientos de comparación múltiple utilizan la idea de un contraste. Considérese el experimento de la concentración de madera dura presentado en la sección 12-2.1. Puesto que la hipótesis H_0: $\tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ fue rechazada, se sabe que algunas de las concentraciones de madera dura producen resistencias a la tensión diferentes que otras, pero, ¿cuáles son las que causan en realidad esta diferencia? Al principio del experimento, podría sospecharse que las concentraciones de madera dura 3 y 4 producen la misma resistencia a la tensión. Esto implica que querría probarse la hipótesis

$$H_0: \mu_3 = \mu_4$$
$$H_1: \mu_3 \neq \mu_4$$

Esta hipótesis podría probarse utilizando una combinación lineal de los totales de los tratamientos, por ejemplo,

$$y_{3\cdot} - y_{4\cdot} \simeq 0$$

Si se hubiera sospechado que el promedio de las concentraciones de madera dura 1 y 3 no difería del promedio de las concentraciones de madera dura 2 y 4, entonces la hipótesis hubiera sido

$$H_0: \mu_1 + \mu_3 = \mu_2 + \mu_4$$
$$H_1: \mu_1 + \mu_3 \neq \mu_2 + \mu_4$$

lo cual implica que la combinación lineal de los totales de los tratamientos es

$$y_{1\cdot} + y_{3\cdot} - y_{2\cdot} - y_{4\cdot} \simeq 0$$

En general, la comparación de las medias de los tratamientos de interés, implicará una combinación lineal de los totales de los tratamientos como

$$c = \sum_{i=1}^{a} c_i y_{i\cdot}$$

con la restricción de que $\sum_{i=1}^{a} c_i = 0$. A estas combinaciones lineales se les llama **contrastes**. La suma de cuadrados de cualquier contraste es

$$SS_c = \frac{\left(\sum_{i=1}^{a} c_i y_{i\cdot}\right)^2}{n\sum_{i=1}^{a} c_i^2} \tag{12-14}$$

que tiene un solo grado de libertad. Si el diseño es desbalanceado, entonces la comparación de las medias de los tratamientos requiere que $\sum_{i=1}^{a} n_i c_i = 0$, y la ecuación 12-14 pasa a ser

$$SS_c = \frac{\left(\sum_{i=1}^{a} c_i y_{i\cdot}\right)^2}{\sum_{i=1}^{a} n_i c_i^2} \tag{12-15}$$

Un contraste se prueba comparando su suma de cuadrados con el cuadrado medio de error. El estadístico resultante tiene una distribución como la F, con 1 y $N - a$ grados de libertad.

Un caso muy especial del procedimiento anterior es el de los **contrastes ortogonales**. Dos contrastes con coeficientes $\{c_i\}$ y $\{d_i\}$ son ortogonales si

$$\sum_{i=1}^{a} c_i d_i = 0$$

o, en el caso de un diseño desbalanceado, si

$$\sum_{i=1}^{a} n_i c_i d_i = 0$$

Para a tratamientos, un conjunto de $a - 1$ contrastes ortogonales harán la partición de la suma de cuadrados debida a los tratamientos en $a - 1$ sumas de cuadrados independientes, con un solo grado de libertad. Por tanto, las pruebas realizadas en contrastes ortogonales son independientes.

Hay muchas maneras de elegir los coeficientes de los contrastes ortogonales para un conjunto de tratamientos. Por lo general, algún elemento del contexto del experimento deberá sugerir cuáles de las comparaciones son de interés. Por ejemplo, si hay $a = 3$ tratamientos, donde el tratamiento 1 es de control y los tratamientos 2 y 3 son niveles reales del factor de interés para el experimentador, entonces los contrastes ortogonales apropiados podrían ser los siguientes:

$$H_0 : -2\mu_1 + \mu_2 + \mu_3 = 0$$
$$H_0 : \qquad \mu_2 - \mu_3 = 0$$

Obsérvese que el contraste 1 con $c_i = -2, 1, 1$ compara el efecto promedio del factor con el control, mientras que el contraste 2 con $d_i = 0, 1, -1$ compara los dos niveles del factor de interés.

Los coeficientes de contraste deben elegirse antes de llevar a cabo el experimento, ya que si estas comparaciones se seleccionan después de examinar los datos, la mayoría de los experimentadores construirían pruebas que compararían las diferencias de las medias más grandes. Estas diferencias más grandes podrían deberse a la presencia de efectos reales, o podrían ser causadas por el error aleatorio. Si los experimentadores eligen siempre las diferencias más grandes para hacer las comparaciones, inflarán el error tipo I de la prueba, ya que es posible que en un porcentaje excepcionalmente alto de las comparaciones seleccionadas, las diferencias observadas se deban al error.

EJEMPLO 12-3

Considérese el experimento de la concentración de madera dura. Hay cuatro niveles de la concentración de madera dura, y los conjuntos posibles de comparaciones entre estas medias y las comparaciones ortogonales asociadas son

$$H_0: \mu_1 + \mu_4 = \mu_2 + \mu_3 \qquad c = y_{1\cdot} - y_{2\cdot} - y_{3\cdot} + y_{4\cdot}$$
$$H_0: \mu_1 + \mu_2 = \mu_3 + \mu_4 \qquad d = -y_{1\cdot} - y_{2\cdot} + y_{3\cdot} + y_{4\cdot}$$
$$H_0: \mu_1 + \mu_3 = \mu_2 + \mu_4 \qquad e = -y_{1\cdot} + y_{2\cdot} - y_{3\cdot} + y_{4\cdot}$$

Obsérvese que las constantes de los contrastes son ortogonales. Utilizando los datos de la tabla 12-1, los valores numéricos de los contrastes y las sumas de cuadrados se encuentran de la siguiente manera.

$$c = 60 - 94 - 102 + 127 = -9 \qquad SS_c = \frac{(-9)^2}{6(4)} = 3.38$$

$$d = -60 - 94 + 102 + 127 = 75 \qquad SS_d = \frac{(75)^2}{6(4)} = 234.38$$

$$e = -60 + 94 - 102 + 127 = 59 \qquad SS_e = \frac{(59)^2}{6(4)} = 145.04$$

Estas sumas de cuadrados de los contrastes hacen una partición completa de la suma de cuadrados de los tratamientos; es decir, $SS_{\text{Tratamientos}} = SS_c + SS_d + SS_e$. Estas pruebas de los contrastes suelen incluirse en los análisis de varianza, como el que se muestra en la tabla 12-7. A partir de este análisis, se concluye que hay diferencias significativas entre las concentraciones de madera dura 3 y 4, y 1 y 2, pero que el promedio de 1 y 4 no difiere del promedio de 2 y 3. Asimismo, el promedio de 1 y 3 difiere del promedio de 2 y 4.

Tabla 12-7 Análisis de varianza de los datos de la resistencia

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0
Concentración de madera dura	382.79	3	127.60	19.61
c (1, 4 vs. 2, 3)	3.38	1	3.38	0.52
d (1, 2 vs. 3, 4)	234.38	1	234.38	36.00
e (1, 3 vs. 2, 4)	145.04	1	145.04	22.28
Error	130.17	20	6.51	
Total	512.96	23		

12-3.3 Método de la diferencia significativa mínima de Fisher

Frecuentemente, los analistas no saben de antemano cómo construir los contrastes ortogonales apropiados, o quizá deseen probar más de $a - 1$ comparaciones usando los mismos datos. Por ejemplo, tal vez un analista quiera comparar todos los pares posibles de medias. La hipótesis nula sería entonces H_0: $\mu_i = \mu_j$, para toda $i \neq j$. Pueden aplicarse varios procedimientos en este problema. Se describe aquí el método de la diferencia significativa mínima (LSD) de Fisher. Para otros métodos y mayores detalles, incluyendo un análisis comparativo de varios métodos, véase Montgomery.

El método de la LSD de Fisher compara todos los pares de medias con las hipótesis nulas H_0: $\mu_i = \mu_j$ (para toda $i \neq j$) utilizando el estadístico t

$$t_0 = \frac{\bar{y}_{i\cdot} - \bar{y}_{j\cdot}}{\sqrt{\dfrac{2MS_E}{n}}}$$

Suponiendo una hipótesis alternativa de dos colas, el par de medias μ_i y μ_j se declararía significativamente diferente si

$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > \text{LSD}$$

donde LSD, la **diferencia significativa mínima**, es

$$\text{LSD} = t_{\alpha/2,\, a(n-1)}\sqrt{\frac{2MS_E}{n}} \tag{12-16}$$

Si los tamaños de las muestras son diferentes en cada tratamiento, entonces la LSD se define como

$$\text{LSD} = t_{\alpha/2,\, N-a}\sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

EJEMPLO 12-4

Se aplicará el método de la LSD de Fisher al experimento de la concentración de madera dura. Hay $a = 4$ medias, $n = 6$, $MS_E = 6.51$ y $t_{0.025,\,20} = 2.086$. Las medias de los tratamientos son

$$\bar{y}_{1\cdot} = 10.00 \text{ psi}$$
$$\bar{y}_{2\cdot} = 15.67 \text{ psi}$$
$$\bar{y}_{3\cdot} = 17.00 \text{ psi}$$
$$\bar{y}_{4\cdot} = 21.17 \text{ psi}$$

El valor de la LSD es $\text{LSD} = t_{0.025,\,20}\sqrt{2MS_E/n} = 2.086\sqrt{2(6.51)/6} = 3.07$. Por lo tanto, cualquier par de promedios de los tratamientos que difiere en más de 3.07 implica que el par de medias de los tratamientos correspondientes son diferentes.

Las comparaciones entre los promedios observados de los tratamientos son las siguientes:

$$
\begin{aligned}
4 \text{ vs. } 1 &= 21.17 - 10.00 = 11.17 > 3.07 \\
4 \text{ vs. } 2 &= 21.17 - 15.67 = 5.50 > 3.07 \\
4 \text{ vs. } 3 &= 21.17 - 17.00 = 4.17 > 3.07 \\
3 \text{ vs. } 1 &= 17.00 - 10.00 = 7.00 > 3.07 \\
3 \text{ vs. } 2 &= 17.00 - 15.67 = 1.33 < 3.07 \\
2 \text{ vs. } 1 &= 15.67 - 10.00 = 5.67 > 3.07
\end{aligned}
$$

A partir de este análisis, se observa que hay diferencias significativas entre todos los pares de medias excepto la 2 y la 3. Esto implica que las concentraciones de madera dura de 10% y 15% producen aproximadamente la misma resistencia a la tensión y que los demás niveles de concentración probados producen resistencias a la tensión diferentes. Con frecuencia es conveniente trazar una gráfica de las medias de los tratamientos, como la de la figura 12-6, con las medias que *no* son diferentes subrayadas. Esta gráfica revela claramente los resultados del experimento y muestra que 20% de madera dura produce la resistencia a la tensión máxima. Obsérvese que estos resultados concuerdan con los del método gráfico simple para comparaciones múltiples descrito en la sección 12-3.1.

El programa Minitab ejecuta varios procedimientos de comparación múltiple. La salida del programa Minitab de la tabla 12-5 muestra el método de la LSD de Fisher bajo el encabezado "Comparaciones por pares de Fisher". El valor crítico reportado es en realidad el valor de $t_{0.025,20} = 2.086$. Minitab aplica el método de la LSD de Fisher calculando intervalos de confianza para todos los pares de medias de los tratamientos utilizando la ecuación 12-13. Los límites de confianza inferior y superior de 95% se muestran en la parte inferior de la tabla. Obsérvese que el único par de medias para los que el intervalo de confianza incluye al cero es para μ_{10} y μ_{15}. Esto implica que μ_{10} y μ_{15} no difieren significativamente, el mismo resultado al que se llegó en el ejemplo 12-4.

En la tabla 12-5 también se proporciona un "Índice de error de familia", igual a 0.192 en este ejemplo. Cuando se prueban todos los pares posibles de medias, la probabilidad de al menos un error tipo I puede ser mucho mayor que para una prueba sencilla. El índice de error de familia puede interpretarse de la siguiente manera. La probabilidad de que no haya errores tipo I en las seis comparaciones es $1 - 0.192 = 0.808$. El índice de error de familia de la tabla 12-5 se basa en la distribución del rango de las medias muestrales. Véase Montgomery para los detalles. Alternativamente, Minitab permite que el usuario especifique un índice de error de familia y después el programa calcula un índice de error de familia para cada comparación.

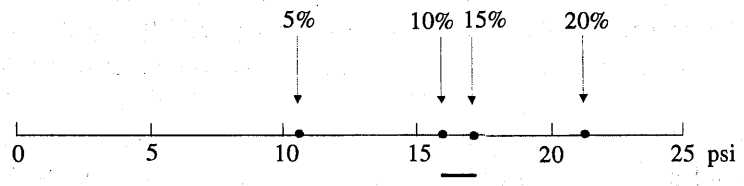


Figura 12-6 Resultados del método de la LSD de Fisher en el ejemplo 12-4.

EJERCICIOS DE LA SECCIÓN 12-3

12-11. Aplique el método gráfico de la sección 12-3.1 para investigar la resistencia a la tensión para los cinco niveles de contenido de algodón del ejercicio 12-1. ¿Cuáles son las conclusiones?

12-12. Aplique el método gráfico de la sección 12-3.1 al experimento de la velocidad de flujo del gas del ejercicio 12-2. ¿Qué conclusiones pueden sacarse acerca de las velocidades de flujo?

12-13. Aplique el método gráfico de la sección 12-3.1 al experimento del ejercicio 12-3. ¿Qué conclusiones pueden sacarse?

12-14. Aplique el método gráfico de la sección 12-3.1 al experimento del ejercicio 12-5. ¿Qué tipos de recubrimientos dan como resultado respuestas medias diferentes?

12-15. Aplique el método gráfico de la sección 12-12-3.1 al experimento del diseño del circuito del ejercicio 12-6. ¿Qué diseños del circuito producen tiempos de respuesta medios diferentes?

12-16. Aplique el método gráfico de la sección 12-3.1 al experimento del radón del ejercicio 12-8. ¿Qué diámetros de los orificios de las regaderas dan como resultado cantidades medias de radón liberado diferentes?

12-17. Aplique la técnica gráfica de la sección 12-3.1 al experimento de los vacíos de aire del ejercicio 12-9. ¿Qué niveles de vacíos de aire dan como resultado resistencias conservadas medias del asfalto diferentes?

12-18. Aplique el método gráfico de la sección 12-3.1 al experimento del material superconductor del ejercicio 12-10. ¿Qué métodos de preparación ocasionan las diferencias en la temperatura media de transición de superconducción?

12-19. Use el método de la LSD de Fisher con $\alpha = 0.05$ para analizar las medias de los cinco niveles diferentes del contenido de algodón del ejercicio 12-1.

12-20. Use el método de la LSD de Fisher con $\alpha = 0.05$ para analizar las medias de las tres velocidades de flujo del ejercicio 12-2.

12-21. Considere el experimento de la velocidad de flujo del gas descrito en el ejercicio 12-2. Suponga que 125 SCCM es una velocidad de flujo estándar. Sugiera un conjunto de contrastes ortogonales que serían apropiados para la comparación de las medias. Haga el análisis usando $\alpha = 0.05$.

12-22. Use el método de la LSD de Fisher con $\alpha = 0.05$ para analizar la resistencia de compresión media de las cuatro técnicas de mezclado del ejercicio 12-3.

12-23. Use el método de la LSD de Fisher para analizar las cinco medias de los tipos de recubrimiento descritos en el ejercicio 12-5. Use $\alpha = 0.01$.

12-24. Use el método de la LSD de Fisher para analizar los tiempos de respuesta media de los tres circuitos descritos en el ejercicio 12-6. Use $\alpha = 0.01$.

12-25. Considere el experimento descrito en el ejercicio 12-5. Suponga que, antes de realizar el experimento, se le pidiera a usted que construyera un conjunto de contrastes ortogonales para este experimento. Proponga un conjunto de contrastes, suponiendo que el recubrimiento tipo 5 es el que está en uso actualmente. ¿Qué conclusiones sacaría usted, usando $\alpha = 0.01$?

12-26. Proponga un conjunto de contrastes ortogonales para comparar las medias del ejercicio 12-6. Analice estos contrastes usando $\alpha = 0.01$ y saque conclusiones.

12-27. Use el método de la LSD de Fisher para analizar las cantidades medias del radón liberado en el experimento descrito en el ejercicio 12-8. Use $\alpha = 0.05$.

12-28. Use el método de la LSD de Fisher en el experimento de los vacíos de aire descrito en el ejercicio 12-9. Usando $\alpha = 0.05$, ¿cuáles medias de los tratamientos son diferentes?

12-29. Aplique el método de la LSD de Fisher en el experimento del material superconductor descrito en el ejercicio 12-10. ¿Qué métodos de preparación difieren si $\alpha = 0.05$?

12-30. Considere el experimento del material superconductor descrito en el ejercicio 12-10. Proponga un conjunto de contrastes ortogonales para este experimento que hubiera sido apropiado antes de que se llevara a cabo realmente.

12-4 MODELO DE EFECTOS ALEATORIOS

En muchas situaciones, el factor de interés tiene un gran número de niveles posibles. El analista se interesa en sacar conclusiones acerca de la población completa de los niveles del factor. Si el experimentador selecciona aleatoriamente a de estos niveles entre la población de los niveles del factor, entonces se dice que el factor es un **factor aleatorio**. Puesto que los niveles del factor que se usan realmente en el experimento se eligieron de manera aleatoria, las conclusiones alcanzadas serán válidas para la población completa de los niveles del factor. Se supondrá que la población de los niveles del factor es de tamaño infinito o bien lo suficientemente grande como para considerarlo infinito.

El modelo lineal estadístico es

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{12-17}$$

donde los efectos de los tratamientos τ_i y los errores ε_{ij} son variables aleatorias independientes. Obsérvese que la estructura del modelo es idéntica a la del caso de efectos fijos, pero los parámetros tienen una interpretación diferente. Si la varianza de los efectos de los tratamientos τ_i es σ_τ^2, entonces, por la independencia la varianza de la respuesta, es

$$V(Y_{ij}) = \sigma_\tau^2 + \sigma^2$$

A las varianzas σ_τ^2 y σ^2 se les llama los **componentes de la varianza**, y al modelo, la ecuación 12-17, se le llama el **modelo de componentes de la varianza** o **modelo de efectos aleatorios**. Para probar hipótesis en este modelo, se supone que los errores ε_{ij} tienen una distribución normal e independiente con media 0 y varianza σ^2, y que los efectos de los tratamientos τ_i tienen una distribución normal e independiente con media cero y varianza σ_τ^2.[1]

Para el modelo de efectos aleatorios, probar la hipótesis de que los efectos de los tratamientos individuales son cero no tiene sentido. Es más apropiado probar una hipótesis acerca de σ_τ^2. Específicamente,

$$H_0 : \sigma_\tau^2 = 0$$
$$H_1 : \sigma_\tau^2 > 0$$

Si $\sigma_\tau^2 = 0$, todos los tratamientos son idénticos; pero si $\sigma_\tau^2 > 0$, entonces hay variabilidad entre los tratamientos.

El análisis de descomposición de varianza de la variabilidad total sigue siendo válido; es decir,

$$SS_T = SS_{\text{Tratamientos}} + SS_E \tag{12-18}$$

[1] El supuesto de que las $\{\tau_i\}$ son variables aleatorias independientes implica que el supuesto usual de $\sum_{i=1}^{a} \tau_i = 0$ del modelo de efectos fijos no se aplica al modelo de efectos aleatorios.

Sin embargo, los valores esperados de los cuadrados medios de los tratamientos y del error presentan ciertas diferencias con respecto al caso de los efectos fijos.

En el modelo de efectos aleatorios para un experimento completamente aleatorizado con un solo factor, el cuadrado medio esperado de los tratamientos es

$$E(MS_{\text{Tratamientos}}) = E\left(\frac{SS_{\text{Tratamientos}}}{a-1}\right) = \sigma^2 + n\sigma_\tau^2 \tag{12-19}$$

y el cuadrado medio esperado del error es

$$E(MS_E) = E\left[\frac{SS_E}{a(n-1)}\right] = \sigma^2 \tag{12-20}$$

El desarrollo de estos cuadrados medios esperados es relativamente simple. Considérese el valor esperado del cuadrado medio de los tratamientos:

$$E(MS_{\text{Tratamientos}}) = \frac{1}{a-1}E(SS_{\text{Tratamientos}}) = \frac{1}{a-1}E\left[\sum_{i=1}^{a}\frac{Y_{i\cdot}^2}{n} - \frac{Y_{\cdot\cdot}^2}{N}\right]$$
$$= \frac{1}{a-1}E\left\{\frac{1}{n}\sum_{i=1}^{a}\left[\sum_{j=1}^{n}(\mu+\tau_i+\varepsilon_{ij})\right]^2 - \frac{1}{N}\left[\sum_{i=1}^{a}\sum_{j=1}^{n}(\mu+\tau_i+\varepsilon_{ij})\right]^2\right\}$$

Si se eleva al cuadrado y se calcula el valor esperado, se encuentra que los términos que incluyen a τ_i^2 son reemplazados por σ_τ^2, ya que $E(t_i) = 0$. Asimismo, los términos que incluyen $\sum_{j=1}^{n}\varepsilon_{ij}^2$, $\sum_{i=1}^{a}\sum_{j=1}^{n}\varepsilon_{ij}^2$ y $\sum_{i=1}^{a}\sum_{j=1}^{n}\tau_i^2$ son reemplazados por $n\sigma^2$, $an\sigma^2$ y $an\sigma_\tau^2$, respectivamente. Por último, todos los términos con productos cruzados que incluyen a τ_i y ε_{ij} tienen expectativa cero porque las variables aleatorias son independientes. Esto lleva a

$$E(MS_{\text{Tratamientos}}) = \frac{1}{a-1}\left[N\mu^2 + N\sigma_\tau^2 + a\sigma^2 - N\mu^2 - n\sigma_\tau^2 - \sigma^2\right]$$

o bien

$$E(MS_{\text{Tratamientos}}) = \sigma^2 - n\sigma_\tau^2$$

como se estableció en la definición. Utilizando un enfoque similar se demuestra con facilidad que $E(MS_E) = \sigma^2$.

Al examinar los cuadrados medios esperados, es claro que tanto MS_E como $MS_{\text{Tratamientos}}$ estiman σ^2 cuando $H_0 : \sigma_\tau^2 = 0$ es verdadera. Además, MS_E y $MS_{\text{Tratamientos}}$ son independientes. Por consiguiente, el cociente

$$F_0 = \frac{MS_{\text{Tratamientos}}}{MS_E} \tag{12-21}$$

es una variable aleatoria F con $a - 1$ y $a(n - 1)$ grados de libertad cuando H_0 es verdadera. La hipótesis nula se rechazaría en el nivel de significación α si el valor calculado del estadístico de la prueba $f_0 > f_{\alpha, a-1, a(n-1)}$.

El procedimiento de cálculo y la construcción de la tabla de análisis de varianza para el modelo de efectos aleatorios son idénticos al caso de efectos fijos. Sin embargo, las conclusiones son muy diferentes, ya que se aplican a la población completa de los tratamientos.

Por lo general, también querrán estimarse los componentes de la varianza (σ^2 y σ_τ^2) del modelo. Al procedimiento que se usará aquí para estimar σ^2 y σ_τ^2 se le llama el **método del análisis de varianza**, porque utiliza la información de la tabla del análisis de varianza. No es necesario el supuesto de normalidad de las observaciones. El procedimiento consiste en igualar los cuadrados medios esperados con sus valores observados en la tabla del análisis de varianza y resolver para los componentes de la varianza. Cuando se igualan los cuadrados medios observados y los esperados del modelo de efectos aleatorios con clasificación simple, se obtiene

$$MS_{\text{Tratamientos}} = \sigma^2 + n\sigma_\tau^2$$

y

$$MS_E = \sigma^2$$

Por lo tanto, los estimadores de los componentes de la varianza son

$$\hat{\sigma}^2 = MS_E \tag{12-22}$$

y

$$\hat{\sigma}_\tau^2 = \frac{MS_{\text{Tratamientos}} - MS_E}{n} \tag{12-23}$$

En ocasiones el método del análisis de varianza produce una estimación negativa de un componente de la varianza. Puesto que los componentes de la varianza por definición son no negativos, una estimación negativa de un componente de la varianza es perturbante. Un curso de acción es aceptar la estimación y usarla como evidencia de que el verdadero valor del componente de la varianza es cero, al suponer que la variación del muestreo llevó a la estimación negativa. Aun cuando este enfoque posee una apelación intuitiva, perturbará las propiedades estadísticas de otras estimaciones. Otra alternativa es estimar de nuevo el componente de la

varianza negativo con un método que produzca siempre estimaciones no negativas. Una posibilidad más es considerar la estimación negativa como evidencia de que el modelo lineal propuesto es incorrecto, lo cual requeriría la revisión del modelo y sus supuestos para encontrar un modelo más apropiado.

EJEMPLO 12-5

En *Design and Analysis of Experiments*, 4a. edición (John Wiley), D.C. Montgomery describe un experimento con un solo factor referente a un modelo de efectos aleatorios en el que una compañía manufacturera de textiles elabora una tela en un gran número de telares. La compañía está interesada en la variabilidad de la resistencia a la tensión de un telar a otro. Para investigar esta variabilidad, un ingeniero de manufactura selecciona cuatro telares al azar y hace cuatro determinaciones de la resistencia en muestras de la tela elegidas al azar de cada telar. Los datos se muestran en la tabla 12-8, y el análisis de varianza se resume en la tabla 12-9.

A partir del análisis de varianza, se concluye que los telares de la planta difieren significativamente en su capacidad para producir una tela de resistencia uniforme. Los componentes de la varianza se estiman con $\hat{\sigma}^2 = 1.90$ y

$$\hat{\sigma}_\tau^2 = \frac{29.73 - 1.90}{4} = 6.96$$

Por lo tanto, la varianza de la resistencia en el proceso de manufactura se estimó por

$$\begin{aligned}\hat{V}(Y_{ij}) &= \sigma_\tau^2 + \hat{\sigma}^2 \\ &= 6.96 + 1.90 \\ &= 8.86\end{aligned}$$

La mayor parte de esta variabilidad es atribuible a las diferencias entre los telares.

Este ejemplo ilustra una importante aplicación del análisis de varianza: el aislamiento de diferentes fuentes de variabilidad en un proceso de manufactura. En los programas de mejoramiento de calidad con frecuencia surgen problemas de variabilidad excesiva en los parámetros

Tabla 12-8 Datos de la resistencia del ejemplo 12-5

	Observaciones					
Telar	1	2	3	4	Total	Promedio
1	98	97	99	96	390	97.5
2	91	90	93	92	366	91.5
3	96	95	97	95	383	95.8
4	95	96	99	98	388	97.0
					1 527	95.45

Tabla 12-9 Análisis de varianza de los datos de la resistencia

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Telares	89.19	3	29.73	15.68	1.88 E-4
Error	22.75	12	1.90		
Total	111.94	15			

o las propiedades funcionales críticas. Por ejemplo, en el caso anterior de la resistencia de la tela, la media del proceso se estima con $\bar{y} = 95.45$ psi, y la desviación estándar del proceso se estima con $\hat{\sigma}_Y = \sqrt{\hat{V}(Y_{ij})} = \sqrt{8.86} = 2.98$ psi. Si la resistencia tiene una distribución aproximadamente normal, entonces la distribución de la resistencia en la producción final, sería parecida a la distribución normal ilustrada en la figura 12-7*a*. Si el límite inferior de las especificaciones (LSL) de la resistencia se localiza en 90 psi, entonces una proporción sustancial de la producción del proceso es **fallida**; es decir, material defectuoso o de desecho que debe venderse como material de segunda calidad. Este material fallido está relacionado directamente con la variabilidad excesiva que resulta de las **diferencias entre los telares**. La variabilidad en el desempeño de los telares podría ser causada por un ajuste incorrecto de la maquinaria, falta de mantenimiento, supervisión inadecuada, operadores sin el entrenamiento adecuado, etc. El ingeniero o gerente responsable del mejoramiento de la calidad debe identificar y eliminar estas fuentes de variabilidad del proceso. Si puede hacerlo, entonces la variabilidad de la resistencia se reducirá en gran medida, quizás a un valor tan bajo como $\hat{\sigma}_Y = \sqrt{\hat{\sigma}^2} = \sqrt{1.90} = 1.38$ psi, como se muestra en la figura 12-7*b*. En este proceso mejorado, la reducción de la variabilidad de la resistencia ha reducido en gran medida la producción fallida, dando como resultado una disminución del costo, calidad más alta, clientes más satisfechos y una posición competitiva fortalecida para la compañía.

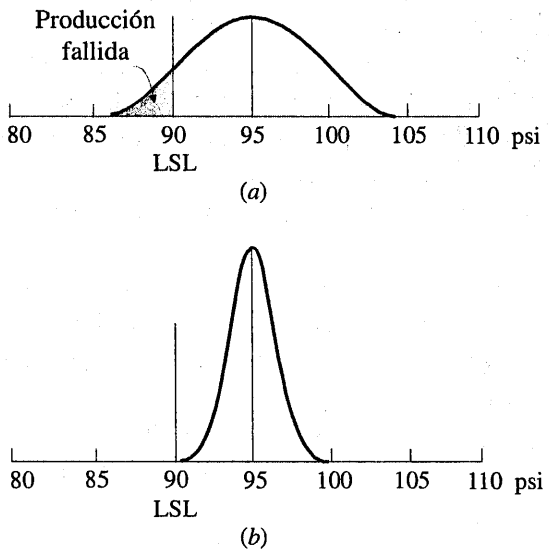


Figura 12-7 Distribución de la resistencia de la tela, ***a***) proceso actual, ***b***) proceso mejorado.

EJERCICIOS DE LA SECCIÓN 12-4

12-31. Una fábrica de textiles tiene un gran número de telares. Se supone que cada telar tiene la misma producción de tela por minuto. Para investigar este supuesto, se eligen cinco telares al azar y se mide su producción en tiempos diferentes. Se obtuvieron los siguientes datos:

Telar	Producción (lb/min)				
1	4.0	4.1	4.2	4.0	4.1
2	3.9	3.8	3.9	4.0	4.0
3	4.1	4.2	4.1	4.0	3.9
4	3.6	3.8	4.0	3.9	3.7
5	3.8	3.6	3.9	3.8	4.0

a) ¿Los telares tienen una producción similar? Use $\alpha = 0.05$.
b) Estime la variabilidad entre los telares.
c) Estime la varianza del error experimental.
d) Analice los residuales de este experimento y verifique la adecuación del modelo.

12-32. En un artículo de *Journal of the Electrochemical Society* (Vol. 139, No. 2, pp. 524-532) se describe un experimento para investigar la deposición de vapor de baja presión del polisilicio. El experimento se realizó en un reactor de alta capacidad en Austin, Texas. El reactor tiene varias posiciones de las obleas, y se seleccionaron al azar cuatro de estas posiciones. La variable de respuesta es la uniformidad del espesor del recubrimiento. Se corrieron tres réplicas del experimento y los datos fueron los siguientes:

Posición de la oblea	Uniformidad		
1	2.76	5.67	4.49
2	1.43	1.70	2.19
3	2.34	1.97	1.47
4	0.94	1.36	1.65

a) ¿Hay alguna diferencia en las posiciones de las obleas? Use $\alpha = 0.05$.
b) Estime la variabilidad debida a la posición de las obleas.
c) Estime el componente del error aleatorio.
d) Analice los residuales de este expermento y comente la adecuación del modelo.

12-33. En un artículo de *Journal of Quality Technology* (Vol. 13, No. 2, pp. 111-114) se describe un experimento que investiga los efectos de cuatro sustancias químicas blanqueadoras sobre la brillantez de la pulpa. Estas cuatro sustancias químicas se seleccionaron al azar de una gran población de agentes blanqueadores potenciales. Los datos son los siguientes.

Sustancia química	Brillantez de la pulpa				
1	77.199	74.466	92.746	76.208	82.876
2	80.522	79.306	81.914	80.346	73.385
3	79.417	78.017	91.596	80.802	80.626
4	78.001	78.358	77.544	77.364	77.386

a) ¿Hay alguna diferencia en los tipos de sustancias químicas? Use $\alpha = 0.05$.
b) Estime la variabilidad debida al tipo de sustancia química.
c) Estime la variabilidad debida al error aleatorio.
d) Analice los residuales de este experimento y comente la adecuación del modelo.

12-34. Considere el experimento de la deposición de vapor del experimento descrito en el ejercicio 12-32.
a) Estime la variabilidad total de la respuesta de uniformidad.
b) ¿Qué cantidad de la variabilidad total de la respuesta de uniformidad se debe a la diferencia entre las posiciones en el reactor?
c) ¿Hasta qué nivel podría reducirse la variabilidad de la respuesta de uniformidad si pudiera eliminarse la variabilidad de una posición a otra en el reactor? ¿Se considera que ésta es una reducción significativa?

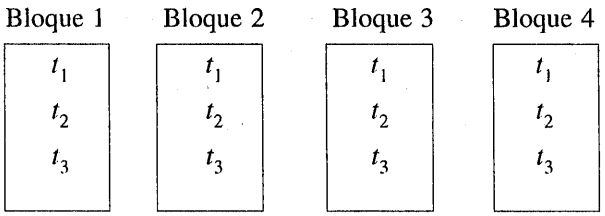


Figura 12-8 Un diseño de bloques completos aleatorizados.

12-5 DISEÑO DE BLOQUES COMPLETOS ALEATORIZADOS

12-5.1 Diseño y análisis estadístico

En muchos problemas de diseño experimental, es necesario diseñar el experimento de tal modo que la variabilidad que surge de un factor perturbador pueda controlarse. Por ejemplo, considérese la situación del ejemplo 9-9, donde se usaron dos métodos diferentes para predecir la resistencia al corte de vigas de placas de acero. Debido a que cada viga tiene (potencialmente) una resistencia diferente, y a que esta variabilidad de la resistencia no era de interés directo, el experimento se diseñó utilizando los dos métodos de prueba en cada viga y después comparando con cero la diferencia promedio en los registros de la resistencia de cada viga usando la prueba t pareada. La prueba t pareada es un procedimiento para comparar las medias de dos tratamientos cuando no es posible hacer todas las corridas experimentales bajo condiciones homogéneas. De manera alternativa, la prueba t pareada puede verse como un método para reducir el ruido ambiental en el experimento separando en bloque (bloqueando) el efecto del **factor perturbador**. El bloque es el factor perturbador, y en este caso, el factor perturbador es la **unidad experimental** en sí, los ejemplares de prueba de las vigas de acero usadas en el experimento.

El diseño de bloques aleatorizados es una ampliación de la prueba t pareada a situaciones en las que el factor de interés tiene más de dos niveles; es decir, deben compararse más de dos tratamientos. Por ejemplo, suponer que podrían usarse tres métodos para evaluar las lecturas de la resistencia de las vigas de placas de acero. Estos métodos pueden considerarse tres tratamientos, por ejemplo t_1, t_2 y t_3. Si se usaran cuatro vigas como unidades experimentales, entonces un **diseño de bloques completos aleatorizados** se vería como se ilustra en la figura 12-8. Al diseño se le llama de bloques completos aleatorizados, porque cada bloque es lo suficientemente grande para contener todos los tratamientos y porque la asignación real de cada uno de los tres tratamientos dentro de cada bloque se hace al azar. Una vez que se ha realizado el experimento, los datos se registran en una tabla, como la que se muestra en la tabla 12-10. Las observaciones de esta tabla, y_{ij}, representan la respuesta obtenida cuando el método i-ésimo se aplica en la viga j-ésima.

Tabla 12-10 Diseño de bloques completos aleatorios

Tratamiento	**Bloque (Viga)**			
(Método)	1	2	3	4
1	y_{11}	y_{12}	y_{13}	y_{14}
2	y_{21}	y_{22}	y_{23}	y_{24}
3	y_{31}	y_{32}	y_{33}	y_{34}

El procedimiento general de un diseño de bloques completos aleatorizados consiste en seleccionar b bloques y correr una réplica completa del experimento en cada bloque. En la tabla 12-11 se muestran los datos que resultan de correr un diseño de bloques completos aleatorizados para investigar un solo factor con a niveles y b bloques. Habrá a observaciones (una por cada nivel del factor) en cada bloque, y el orden en que se corren estas observaciones se asigna aleatoriamente dentro del bloque.

Se describirá ahora el análisis estadístico de un diseño de bloques completos aleatorizados. Suponer que el interés se centra en un solo factor con a niveles y que el experimento se corre en b bloques. Las observaciones pueden representarse con el modelo estadístico lineal

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{12-24}$$

donde μ es la media global, τ_i es el efecto del tratamiento i-ésimo, β_j es el efecto del bloque j-ésimo, y ε_{ij} es el término del error aleatorio, el cual se supone que tiene una distribución normal e independiente con media cero y varianza σ^2. Los tratamientos y los bloques se considerarán inicialmente como factores fijos. Además, los efectos de los tratamientos y los bloques se definen como desviaciones de la media global, por lo que $\sum_{i=1}^{a} \tau_i = 0$ y $\sum_{j=1}^{b} \beta_j = 0$. Se supondrá asimismo que no hay interacción entre los tratamientos y los bloques. Es decir, el efecto del tratamiento i-ésimo es el mismo, independientemente del bloque (o bloques) en el que se pruebe. El interés se encuentra en probar la igualdad de los efectos de los tratamientos. Es decir

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$$
$$H_1 : \tau_i \neq 0 \text{ para al menos una } i$$

Probar la hipótesis de que todos los efectos de los tratamientos τ_i son iguales a cero, es equivalente a probar la hipótesis de que las medias de los tratamientos son iguales. Para ver esto, obsérvese que la media del tratamiento i-ésimo es μ_i, definida como

Tabla 12-11 Diseño de bloques completos aleatorizados con a tratamientos y b bloques

	Bloques					
Tratamientos	1	2	...	b	Totales	Promedios
1	y_{11}	y_{12}	...	y_{1b}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	...	y_{2b}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{ab}	$y_{a\cdot}$	$\bar{y}_{a\cdot}$
Totales	$y_{\cdot 1}$	$y_{\cdot 2}$	...	$y_{\cdot b}$	$y_{\cdot\cdot}$	
Promedios	$\bar{y}_{\cdot 1}$	$\bar{y}_{\cdot 2}$	...	$\bar{y}_{\cdot b}$		$\bar{y}_{\cdot\cdot}$

$$\begin{aligned}
\mu_i &= E\left(\frac{\sum_{j=1}^{b} Y_{ij}}{b}\right) \\
&= \frac{1}{b}\sum_{j=1}^{b} E(y_{ij}) \\
&= \frac{1}{b}\sum_{j=1}^{b} E(\mu + \tau_i + \beta_j + \varepsilon_{ij}) \\
&= \frac{1}{b}\sum_{j=1}^{b} (\mu + \tau_i + \beta_j) \\
&= \mu + \tau_i + \frac{1}{b}\sum_{j=1}^{b} \beta_j
\end{aligned}$$

y como $\sum_{j=1}^{b} \beta_j = 0$, se tiene que la media del tratamiento i-ésimo se define como

$$\mu_i = \mu + \tau_i, \quad i = 1, 2, ..., a$$

Por lo tanto, probar la hipótesis de que las medias de los a tratamientos son iguales, es equivalente a probar que todos los efectos de los tratamientos τ_i son iguales a cero.

El análisis de varianza puede extenderse al diseño de bloques completos aleatorizados. En el procedimiento se utiliza la suma de cuadrados identidad que hace la partición de la suma total de cuadrados en tres componentes.

La **suma de cuadrados identidad para el diseño de bloques completos aleatorizados** es

$$\begin{aligned}
\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 &= b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 \\
&\quad + \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..})^2 \qquad (12\text{-}25)
\end{aligned}$$

La demostración de esta identidad se obtiene desarrollando el miembro izquierdo de la ecuación 12-25 como sigue:

$$\begin{aligned}
\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{..})^2 &= \sum_{i=1}^{a}\sum_{j=1}^{b}[(\bar{y}_{i.}-\bar{y}_{..}) \\
&\quad +(\bar{y}_{.j}-\bar{y}_{..})+(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})]^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})^2+\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{.i}-\bar{y}_{..})^2 \\
&\quad +\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})^2 \\
&\quad +2\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{.j}+\bar{y}_{..}) \\
&\quad +2\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..}) \\
&\quad +2\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{.j}+\bar{y}_{..})(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})
\end{aligned}$$

Los tres términos de los productos cruzados de esta última expresión son cero. Por lo tanto, $\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij}-\bar{y}_{..})^2$ se reduce a la identidad de la ecuación 12-25.

La suma de cuadrados identidad puede representarse simbólicamente como

$$SS_T = SS_{\text{Tratamientos}} + SS_{\text{Bloques}} + SS_E$$

donde

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{..})^2 = \text{suma total de cuadrados}$$

$$SS_{\text{Tratamientos}} = b\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})^2 = \text{suma de cuadrados de los tratamientos}$$

$$SS_{\text{Bloques}} = a\sum_{j=1}^{b}(\bar{y}_{.j}-\bar{y}_{..})^2 = \text{suma de cuadrados de los bloques}$$

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij}-\bar{y}_{.j}-\bar{y}_{i.}+\bar{y}_{..})^2 = \text{suma de cuadrados del error}$$

Además, la descomposición de los grados de libertad que corresponden a estas sumas de cuadrados es

$$ab-1=(a-1)+(b-1)+(a-1)(b-1)$$

Para el diseño de bloques aleatorizados, los cuadrados medios apropiados son

$$MS_{\text{Tratamientos}} = \frac{SS_{\text{Tratamientos}}}{a-1}$$

$$MS_{\text{Bloques}} = \frac{SS_{\text{Bloques}}}{b-1}$$

$$MS_E = \frac{SS_E}{(a-1)(b-1)}$$

Puede demostrarse que los valores esperados de estos cuadrados medios son los siguientes:

$$E(MS_{\text{Tratamientos}}) = \sigma^2 + \frac{b\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E(MS_{\text{Bloques}}) = \sigma^2 + \frac{a\sum_{j=1}^{j}\beta_j^2}{b-1}$$

$$E(MS_E) = \sigma^2$$

Por lo tanto, si la hipótesis nula H_0 es verdadera de tal modo que los efectos de los tratamientos $\tau_i = 0$, entonces $MS_{\text{Tratamientos}}$ es un estimador insesgado de σ^2, mientras que si H_0 es falsa, $MS_{\text{Tratamientos}}$ sobrestima σ^2. El cuadrado medio de error siempre es una estimación insesgada de σ^2. Para probar la hipótesis nula de que todos los efectos de los tratamientos son cero, se calcula el cociente

$$F_0 = \frac{MS_{\text{Tratamientos}}}{MS_E} \tag{12-26}$$

que tiene una distribución F con $a - 1$ y $(a - 1)(b - 1)$ grados de libertad, si la hipótesis nula es verdadera. La hipótesis nula se rechazaría con el nivel de significación α si el valor calculado del estadístico de la prueba de la ecuación 12-26

$$f_0 > f_{\alpha,\, a-1,\, (a-1)(b-1)}$$

En la práctica, se calculan SS_T, $SS_{\text{Tratamientos}}$ y SS_{Bloques}, y después se obtiene la suma de cuadrados del error SS_E por sustracción. Las fórmulas apropiadas para realizar los cálculos son las siguientes:

Definición

Las fórmulas para calcular las sumas de cuadrados en el análisis de varianza de un diseño de bloques completos aleatorizados son

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab} \tag{12-27}$$

$$SS_{\text{Tratamientos}} = \frac{1}{b} \sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{ab} \tag{12-28}$$

$$SS_{\text{Bloques}} = \frac{1}{a} \sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{ab} \tag{12-29}$$

y

$$SS_E = SS_T - SS_{\text{Tratamientos}} - SS_{\text{Bloques}} \tag{12-30}$$

Los cálculos suelen presentarse en una tabla de análisis de varianza, como la que se muestra en la tabla 12-12. En general, se usará un paquete de software de computadora para realizar el análisis de varianza del diseño de bloques completos aleatorizados.

Tabla 12-12 Análisis de varianza de un diseño de bloques completos aleatorios

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$SS_{\text{Tratamientos}}$	$a-1$	$\frac{SS_{\text{Tratamientos}}}{a-1}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
Bloques	SS_{Bloques}	$b-1$	$\frac{SS_{\text{Bloques}}}{b-1}$	
Error	SS_E (por sustracción)	$(a-1)(b-1)$	$\frac{SS_E}{(a-1)(b-1)}$	
Total	SS_T	$ab-1$		

EJEMPLO 12-6

Se realizó un experimento para determinar el efecto de cuatro sustancias químicas diferentes sobre la resistencia de una tela. Estas sustancias químicas se usan como parte del proceso de acabado del planchado permanente. Se seleccionaron cinco muestras de tela, y se corrió un diseño de bloques completos aleatorizados para probar cada tipo de sustancia química sobre cada muestra de tela en orden aleatorio. En la tabla 12-13 se muestran los datos. Se probarán las diferencias de las medias utilizando el análisis de varianza con $\alpha = 0.01$.

Tabla 12-13 Datos de la resistencia de la tela: diseño de bloques completos aleatorios

Tipo de sustancia química	Muestra de tela 1	2	3	4	5	Totales de tratamiento $y_{i\cdot}$	Promedios de tratamiento $\bar{y}_{i\cdot}$
1	1.3	1.6	0.5	1.2	1.1	5.7	1.14
2	2.2	2.4	0.4	2.0	1.8	8.8	1.76
3	1.8	1.7	0.6	1.5	1.3	6.9	1.38
4	3.9	4.4	2.0	4.1	3.4	17.8	3.56
Totales de los bloques $y_{\cdot j}$	9.2	10.1	3.5	8.8	7.6	39.2($y_{\cdot\cdot}$)	
Promedios de los bloques $\bar{y}_{\cdot j}$	2.30	2.53	0.88	2.20	1.90		1.96($\bar{y}_{\cdot\cdot}$)

Las sumas de cuadrados del análisis de varianza se calculan como sigue:

$$SS_T = \sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{ab}$$

$$= (1.3)^2 + (1.6)^2 + \cdots + (3.4)^2 - \frac{(39.2)^2}{20} = 25.69$$

$$SS_{\text{Tratamientos}} = \sum_{i=1}^{4} \frac{y_{i\cdot}^2}{b} - \frac{y_{\cdot\cdot}^2}{ab}$$

$$= \frac{(5.7)^2 + (8.8)^2 + (6.9)^2 + (17.8)^2}{5} - \frac{(39.2)^2}{20} = 18.04$$

$$SS_{\text{Bloques}} = \sum_{j=1}^{5} \frac{y_{\cdot j}^2}{a} - \frac{y_{\cdot\cdot}^2}{ab}$$

$$= \frac{(9.2)^2 + (10.1)^2 + (3.5)^2 + (8.8)^2 + (7.6)^2}{4} - \frac{(39.2)^2}{20} = 6.69$$

$$SS_E = SS_T - SS_{\text{Bloques}} - SS_{\text{Tratamientos}}$$

$$= 25.69 - 6.69 - 18.04 = 0.96$$

El análisis de varianza se resume en la tabla 12-14. Puesto que $f_0 = 75.13 > f_{0.01,\,3,\,12} = 5.95$ (el valor P es 4.79×10^{-8}) se concluye que hay una diferencia significativa en los tipos de sustancias químicas en lo que se refiere a su efecto sobre la resistencia promedio de la tela.

¿Cuándo es necesario hacer bloques?

Suponga que un experimento se realiza como un diseño de bloques aleatorizados, cuando en realidad no era necesaria la separación en bloques. Hay ab observaciones y $(a-1)(b-1)$ grados de libertad para el error. Si el experimento se hubiera corrido como un diseño de bloques completos aleatorizados con un solo factor con b réplicas, se habrían tenido $a(b-1)$ grados de

Tabla 12-14 Análisis de varianza del experimento de bloques completos aleatorizados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado Medio	f_0	Valor P
Tipos de sustancias químicas (tratamientos)	18.04	3	6.01	75.13	4.79 E-8
Muestras de tela (bloques)	6.69	4	1.67		
Error	0.96	12	0.08		
Total	25.69	19			

libertad para el error. Por lo tanto, la separación en bloques ha costado $a(b-1)-(a-1)(b-1) = b-1$ grados de libertad para el error. Por tanto, puesto que la pérdida en los grados de libertad para el error por lo general es pequeña, si existe una oportunidad razonable de que los efectos de bloque sean importantes, el experimentador deberá usar el diseño de bloques aleatorizados.

Por ejemplo, considérese el experimento que se describió en el ejemplo 12-6 como un experimento con un solo factor sin separación en bloques. Se tendrían entonces 16 grados de libertad para el error. En el diseño de bloques completos aleatorizados hay 12 grados de libertad para el error. Por lo tanto, el diseño de bloques ha costado tan sólo 4 grados de libertad, pérdida que es muy pequeña considerando la posible ganancia de información que se lograría si los efectos de bloque en realidad son importantes. El efecto de bloque del ejemplo 12-6 es grande, y si no se hubiera utilizado este diseño, SS_{Bloques} se habría incluido en la suma de cuadrados del error, en el análisis completamente aleatorizado. Esto habría dado como resultado un valor de MS_E mucho más grande, haciendo más difícil detectar las diferencias de los tratamientos. Como regla general, cuando se tengan dudas acerca de la importancia de los efectos de bloque, el experimentador deberá hacer este diseño y apostar a que el efecto de bloque existe. Si el experimentador se equivoca, la ligera pérdida en los grados de libertad del error tendrán un efecto omisible, a menos que el número de grados de libertad sea muy pequeño.

Solución de computadora

En la tabla 12-15 se presenta la salida de computadora del programa Minitab para el diseño de bloques completos aleatorizados del ejemplo 12-6. Se utilizó el menú del análisis de varianza para diseños balanceados para resolver este problema. Los resultados concuerdan bastante bien con los cálculos manuales de la tabla 12-14. Obsérvese que Minitab calcula un estadístico F para los bloques (las muestras de tela). La validez de este cociente como estadístico de la prueba para la hipótesis nula de que no hay efectos de bloque es dudosa, ya que los bloques representan una **restricción sobre la aleatorización**; es decir, la aleatorización sólo se ha hecho al interior de los bloques. Si los bloques no se eligen al azar, o si no se corren en orden aleatorio, entonces el cociente F de los bloques quizá no proporcione información confiable acerca de los efectos de bloque. Para una revisión más amplia, véase Montgomery (capítulo 5).

12-5.2 Pruebas sobre medias de tratamientos individuales

Cuando el análisis de varianza indica que existe una diferencia entre las medias de los tratamientos, quizá sea necesario realizar algunas pruebas de seguimiento para aislar las diferencias específicas.

Tabla 12-15 Análisis de varianza del programa Minitab para el diseño de bloques completos aleatorizados del ejemplo 12-6

Análisis de varianza (diseños balanceados)

Factor	Tipo	Niveles	Valores				
Sustancia química	fijo	4	1	2	3	4	
Muestra de tela	fijo	5	1	2	3	4	5

Análisis de varianza de la resistencia

Fuente	DF	SS	MS	F	P
Sustancia química	3	18.0440	6.0147	75.89	0.000
Muestra de tela	4	6.6930	1.6733	21.11	0.000
Error	12	0.9510	0.0792		
Total	19	25.6880			

Prueba F con denominador: Error
Denominador MS = 0.079250 con 12 grados de libertad

Numerador	DF	MS	F	P
Sustancia química	3	6.015	75.89	0.000
Muestra de tela	4	1.673	21.11	0.000

Cualquier método de comparación múltiple, como el método gráfico descrito en la sección 12-3.1 o el método de la LSD de Fisher, podría utilizarse a este fin.

Se ilustrará el método de la LSD de Fisher. Los promedios de los cuatro tipos de sustancias químicas del ejemplo 12-6 son:

$$\bar{y}_{1\cdot} = 1.14$$

$$\bar{y}_{2\cdot} = 1.76$$

$$\bar{y}_{3\cdot} = 1.38$$

$$\bar{y}_{4\cdot} = 3.56$$

El promedio de cada tratamiento usa $b = 5$ observaciones (una de cada bloque). Se usará $\alpha = 0.05$, por lo que $t_{0.025,\,12} = 2.179$. Por lo tanto, el valor de la LSD es

$$\begin{aligned} \text{LSD} &= t_{0.025,\,12}\sqrt{\frac{2MS_E}{b}} \\ &= 2.179\sqrt{\frac{2(0.08)}{5}} \\ &= 0.39 \end{aligned}$$

Cualquier par de promedios de los tratamientos que difiera en 0.39 o más, indica que ese par de medias de los tratamientos presenta una diferencia significativa. Las comparaciones se muestran a continuación:

$$4 \text{ vs. } 1 = \bar{y}_{4\cdot} - \bar{y}_{1\cdot} = 3.56 - 1.14 = 2.42 > 0.39$$
$$4 \text{ vs. } 3 = \bar{y}_{4\cdot} - \bar{y}_{3\cdot} = 3.56 - 1.38 = 2.18 > 0.39$$
$$4 \text{ vs. } 2 = \bar{y}_{4\cdot} - \bar{y}_{2\cdot} = 3.56 - 1.76 = 1.80 > 0.39$$
$$2 \text{ vs. } 1 = \bar{y}_{2\cdot} - \bar{y}_{1\cdot} = 1.76 - 1.14 = 0.62 > 0.39$$
$$2 \text{ vs. } 3 = \bar{y}_{2\cdot} - \bar{y}_{3\cdot} = 1.76 - 1.38 = 0.38 < 0.39$$
$$3 \text{ vs. } 1 = \bar{y}_{3\cdot} - \bar{y}_{1\cdot} = 1.38 - 1.14 = 0.24 < 0.39$$

En la figura 12-9 se presenta la ilustración gráfica de los resultados. Los pares de medias subrayados no son diferentes. El procedimiento de la LSD indica que la sustancia química tipo 4 resulta en resistencias significativamente diferentes que los otros tres tipos. Los químicos tipo 2 y 3 no difieren, y los tipos 1 y 3 no difieren. Puede haber una ligera diferencia en la resistencia entre los tipos 1 y 2.

12-5.3 Análisis de residuales y verificación del modelo

En cualquier experimento diseñado, siempre es importante examinar los residuales y verificar la transgresión de los supuestos básicos que pudiera invalidar los resultados. Como de costumbre, los residuales del diseño de bloques completos aleatorizados, son tan sólo la diferencia entre los valores observados y estimados (o ajustados) del modelo estadístico; es decir,

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$

y los valores ajustados son

$$\hat{y}_{ij} = \bar{y}_{i\cdot} + \bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot} \tag{12-31}$$

El valor ajustado representa la estimación de la respuesta media cuando el tratamiento i-ésimo se aplica en el bloque j-ésimo. Los residuales del experimento del tipo de sustancia química se muestran en la tabla 12-16.

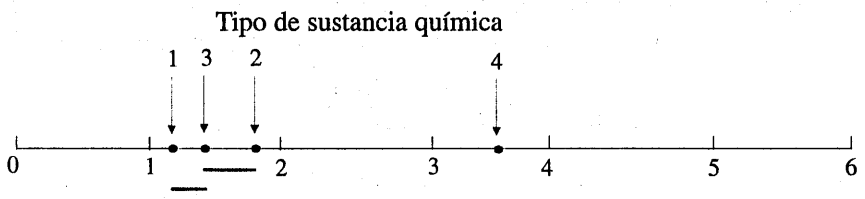


Figura 12-9 Resultados del método de la LSD de Fisher.

Tabla 12-16 Residuales del diseño de bloques completos aleatorizados

Tipo de sustancia química	Muestra de tela 1	2	3	4	5
1	–0.18	–0.10	0.44	–0.18	0.02
2	0.10	0.08	–0.28	0.00	0.10
3	0.08	–0.24	0.30	–0.12	–0.02
4	0.00	0.28	–0.48	0.30	–0.10

En las figuras 12-10, 12-11, 12-12 y 12-13 se presentan las gráficas de los residuales importantes para el experimento. Estas gráficas de los residuales por lo general se hacen con paquetes de software de computadora. Hay indicios de que la muestra (bloque) de tela 3, tiene una variabilidad mayor en la resistencia cuando se trata con las cuatro sustancias químicas que las otras muestras. La sustancia química tipo 4, que produce la resistencia más alta, también tiene una variabilidad un tanto mayor en la resistencia. Pueden ser necesarios experimentos subsecuentes para confirmar estos resultados si son potencialmente importantes.

12-5.4 Diseño de bloques completos aleatorizados con factores aleatorios

En las secciones precedentes, se ha supuesto que los tratamientos y los bloques son factores fijos. En muchos diseños de bloques completos aleatorizados estos supuestos pueden ser demasiado restrictivos. Por ejemplo, en el experimento del tipo de sustancia química, en el ejemplo 12-6, quizá las muestras de tela quisieran considerarse como una muestra aleatoria del material al que pueden aplicarse las sustancias químicas, para que las conclusiones del experimento puedan extenderse a la población completa del material.

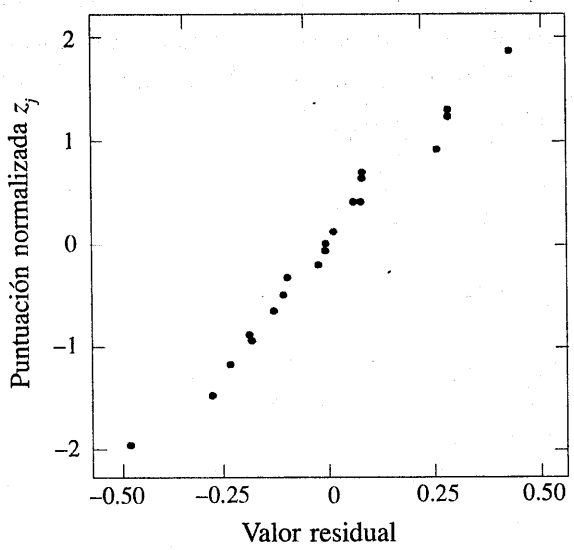


Figura 12-10 Gráfica de probabilidad normal de los residuales del diseño de bloques completos aleatorizados.

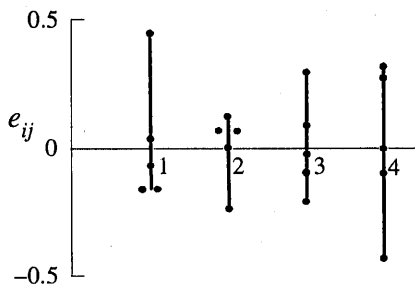


Figura 12-11 Residuales por tratamiento.

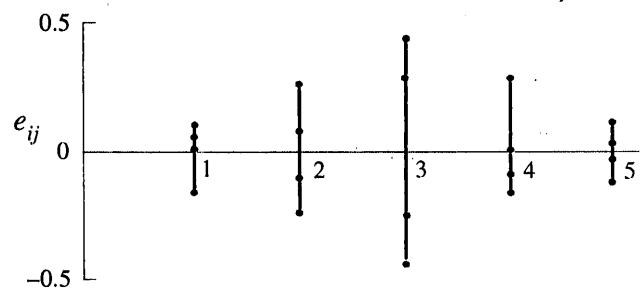


Figura 12-12 Residuales por bloque.

Resulta que si los tratamientos o los bloques (o ambos) son efectos aleatorios, la prueba F en el análisis de varianza se sigue formando como

$$F_0 = \frac{MS_{\text{Tratamientos}}}{MS_E}$$

Esto puede demostrarse utilizando los métodos presentados con anterioridad en este capítulo para evaluar los cuadrados medios esperados. Si los tratamientos son aleatorios, entonces los efectos de los tratamientos τ_i se consideran variables aleatorias con una distribución normal e independiente con media cero y varianza σ_τ^2. La hipótesis nula de cero efectos de los tratamientos es

$$H_0 : \sigma_\tau^2 = 0$$
$$H_1 : \sigma_\tau^2 > 0$$

Cuando tanto los tratamientos como los bloques son aleatorios, entonces se supone también que los efectos de los bloques β_j son variables aleatorias con una distribución normal e independiente con media cero y varianza σ_β^2. En este caso los valores esperados de los cuadrados medios de los tratamientos, los bloques y el error son

$$E(MS_{\text{Tratamientos}}) = \sigma^2 + b\sigma_\tau^2$$
$$E(MS_{\text{Bloques}}) = \sigma^2 + a\sigma_\beta^2$$
$$E(MS_E) = \sigma^2$$

Las estimaciones insesgadas de los componentes de la varianza son

$$\hat{\sigma}^2 = MS_E$$
$$\hat{\sigma}_\tau^2 = \frac{MS_{\text{Tratamientos}} - MS_E}{b}$$
$$\hat{\sigma}_\beta^2 = \frac{MS_{\text{Bloques}} - MS_E}{a}$$

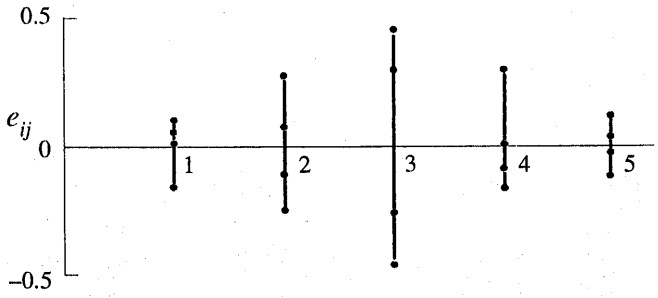


Figura 12-13 Residuales contra $\hat{y}_{ij}$.

EJERCICIOS DE LA SECCIÓN 12-5

12-35. En "El efecto del diseño de boquillas en la estabilidad y el desempeño de surtidores de agua turbulenta" (*Fire Safety Journal*, Vol. 4), C. Theobald describe un experimento en el que se determinó una medición de la forma de varios tipos diferentes de boquillas en diferentes niveles de la velocidad del flujo de salida del surtidor. El interés en este experimento se enfoca principalmente en el tipo de boquilla, y la velocidad es un factor perturbador. Los datos son los siguientes:

Tipo de boquilla	Velocidad del flujo de salida del surtidor (m/s) 11.73	14.37	16.59	20.43	23.46	28.74
1	0.78	0.80	0.81	0.75	0.77	0.78
2	0.85	0.85	0.92	0.86	0.81	0.83
3	0.93	0.92	0.95	0.89	0.89	0.83
4	1.14	0.97	0.98	0.88	0.86	0.83
5	0.97	0.86	0.78	0.76	0.76	0.75

a) ¿El tipo de boquilla afecta la medición de la forma? Compare las boquillas utilizando diagramas de caja y el análisis de varianza.

b) Use el método gráfico de la sección 12-3.1 para determinar diferencias específicas entre las boquillas. ¿Una gráfica del promedio (o desviación estándar) de las mediciones de la forma contra el tipo de boquilla ayuda en las conclusiones?

c) Analice los residuales de este experimento.

12-36. En *Design and Analysis of Experiments*, 4a. edición (John Wiley & Sons), D.C. Montgomery describe un experimento para determinar el efecto de cuatro tipos diferentes de puntas en un probador de dureza, sobre la dureza observada de una aleación metálica. Se obtuvieron cuatro ejemplares de prueba de la aleación, y cada punta se probó una vez en cada ejemplar, produciendo los siguientes datos:

Tipo de punta	Ejemplar de prueba 1	2	3	4
1	9.3	9.4	9.6	10.0
2	9.4	9.3	9.8	9.9
3	9.2	9.4	9.5	9.7
4	9.7	9.6	10.0	10.2

a) ¿Hay alguna diferencia en las mediciones de la dureza entre las puntas?

b) Use el método gráfico de la sección 12-3.1 para investigar diferencias específicas entre las puntas.

c) Analice los residuales de este experimento.

12-37. En un artículo en *American Industrial Hygiene Association Journal* (Vol. 37, pp. 418-422) se describe una prueba de campo para detectar la presencia de arsénico en muestras de orina. La prueba se propuso para aplicarla entre trabajadores forestales debido al uso creciente de arsénicos orgánicos en esa industria. En el experimento se comparó la prueba, aplicada por un practicante

y por un instructor experimentado, con el análisis en un laboratorio remoto. Se seleccionaron cuatro sujetos para someterlos a la prueba y se les consideró como bloques. La variable de respuesta es el contenido de arsénico (en ppm) en la orina de los sujetos. Los datos son los siguientes:

	Sujeto			
Prueba	1	2	3	4
Practicante	0.05	0.05	0.04	0.15
Instructor	0.05	0.05	0.04	0.17
Laboratorio	0.04	0.04	0.03	0.10

a) ¿Hay alguna diferencia en el procedimiento de prueba del arsénico?

b) Analice los residuales de este experimento.

12-38. En un artículo de *Food Technology Journal* (Vol. 10, pp. 39-42) se describe un estudio sobre el contenido de protopectina de jitomates durante el almacenaje. Se seleccionaron cuatro tiempos de almacenaje y se analizaron muestras de nueve lotes de jitomates. El contenido de protopectina (expresado como la fracción soluble de ácido clorhídrico mg/kg) se presenta en la tabla de abajo.

a) Los investigadores de este estudio plantearon la hipótesis de que el contenido medio de protopectina sería diferente en tiempos de almacenaje diferentes. ¿Es posible confirmar esta hipótesis con una prueba estadística usando $\alpha = 0.05$?

b) Encuentre el valor P para la prueba del inciso a).

c) ¿Qué tiempos de almacenaje específicos son diferentes? ¿Estaría usted de acuerdo con la afirmación de que el contenido de protopectina disminuye cuando el tiempo de almacenaje se incrementa?

d) Analice los residuales de este experimento.

12-39. Se realizó un experimento para investigar la dispersión de corriente en la cercanía de una micra de un dispositivo SOS MOSFETS. La finalidad del experimento fue investigar cómo varía la dispersión de corriente cuando cambia la longitud del canal. Se seleccionaron cuatro longitudes del canal. Para cada longitud del canal, se usaron también cinco amplitudes diferentes, y la amplitud deberá considerarse un factor perturbador. Los datos son los siguientes:

	Amplitud				
Longitud del canal	1	2	3	4	5
1	0.7	0.8	0.8	0.9	1.0
2	0.8	0.8	0.9	0.9	1.0
3	0.9	1.0	1.7	2.0	4.0
4	1.0	1.5	2.0	3.0	20.0

a) Pruebe la hipótesis de que el voltaje de dispersión promedio no depende de la longitud del canal, utilizando $\alpha = 0.05$.

b) Analice los residuales de este experimento. Comente las gráficas de los residuales.

12-40. Considere el experimento del voltaje de dispersión descrito en el ejercicio 12-39. El voltaje de dispersión observado para la longitud del canal 4 y la amplitud 5 se registró incorrectamente. La observación correcta es 4.0. Analice los datos corregidos de este experimento. ¿Hay evidencia para concluir que el voltaje de dispersión promedio se incrementa con la longitud del canal?

	Lote								
Tiempo de almacenaje	1	2	3	4	5	6	7	8	9
0 días	1694.0	989.0	917.3	346.1	1260.0	965.6	1123.0	1106.0	1116.0
7 días	1802.0	1074.0	278.8	1375.0	544.0	672.2	818.0	406.8	461.6
14 días	1568.0	646.2	1820.0	1150.0	983.7	395.3	422.3	420.0	409.5
21 días	415.5	845.4	377.6	279.4	447.8	272.1	394.1	356.4	351.2

12-6 DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA EN EXPERIMENTOS CON UN SOLO FACTOR

En cualquier problema de diseño experimental es importante la elección del tamaño de la muestra o del número de réplicas que deberán usarse. Pueden usarse las **curvas de operación característica** como guía para hacer esta selección. Recuérdese que la curva de operación característica es una gráfica de la probabilidad de un error tipo II (β) para varios tamaños de la muestra contra una medida de la diferencia de las medias que es importante detectar. Por tanto, si el experimentador conoce la magnitud de la diferencia de las medias que es de importancia potencial, las curvas de operación característica pueden usarse para determinar cuántas repeticiones se necesitan a fin de alcanzar la sensibilidad adecuada.

12-6.1 Caso con efectos fijos

Se considera primero la determinación del tamaño de la muestra en un modelo con efectos fijos para el caso del mismo tamaño de la muestra en cada tratamiento. La potencia ($1 - \beta$) de la prueba es

$$\begin{aligned} 1 - \beta &= P\{\text{Se rechaza } H_0 \mid H_0 \text{ es falsa}\} \\ &= P\{F_0 > f_{\alpha,\, a-1,\, a(n-1)} \mid H_0 \text{ es falsa}\} \end{aligned} \tag{12-32}$$

Para evaluar este enunciado de probabilidad es necesario conocer la distribución del estadístico de la prueba, F_0, si la hipótesis nula es falsa. Puede demostrarse que, si H_0 es falsa, el estadístico $F_0 = MS_{\text{Tratamientos}} / MS_E$ tiene una distribución como la de una **variable aleatoria F no central**, con $a - 1$ y $a(n - 1)$ grados de libertad y un parámetro de no centralidad δ. Si $\delta = 0$, entonces la distribución F no central se convierte en la distribución F *central* o común.

Las curvas de operación característica dadas en la carta VII del apéndice se usan para evaluar β definido en la ecuación 12-32. En estas curvas se grafica β contra un parámetro Φ, donde

$$\Phi^2 = \frac{n\sum_{i=1}^{a} \tau_i^2}{a\sigma^2} \tag{12-33}$$

El parámetro Φ^2 está relacionado con el parámetro de no centralidad δ. Se dispone de curvas para $\alpha = 0.05$ y $\alpha = 0.01$ y para varios valores de los grados de libertad del numerador y el denominador. En un diseño completamente aleatorizado, el símbolo n de la ecuación 12-33 es el número de réplicas. En un diseño de bloques completos aleatorizados, n se sustituye con el número de bloques.

Cuando se usan las curvas de operación, debe definirse la diferencia de las medias que quiere detectarse en términos de $\sum_{i=1}^{a} \tau_i^2$. Asimismo, por lo general se conoce la varianza del error σ^2. En tales casos, deben elegirse los cocientes de $\sum_{i=1}^{a} \tau_i^2 / \sigma^2$ que quieren detectarse. De manera alternativa, si se cuenta con una estimación de σ^2, puede sustituirse σ^2 con esta estimación. Por ejemplo, si el interés estuviera en la sensibilidad de un experimento que ya se ha realizado, podría usarse MS_E como estimación de σ^2.

EJEMPLO 12-7

Suponga que se están comparando cinco medias en un experimento completamente aleatorizado con $\alpha = 0.01$. El experimentador querría saber cuántas réplicas correr si es importante rechazar H_0 con una probabilidad de al menos 0.90 si $\sum_{i=1}^{5} \tau_i^2 / \sigma^2 = 5.0$. En este caso, el parámetro Φ^2 es,

$$\Phi^2 = \frac{n\sum_{i=1}^{a} \tau_i^2}{a\sigma^2} = \frac{n}{5}(5) = n$$

y para la curva de operación característica para $a - 1 = 5 - 1 = 4$, y $a(n - 1) = 5(n - 1)$ grados de libertad del error se consulta la carta VII del apéndice. Como primera conjetura, se prueba con $n = 4$ réplicas. Con este valor se obtiene $\Phi^2 = 4$, $\Phi = 2$, y $5(3) = 15$ grados de libertad del error. Por consiguiente, por la carta VII, se encuentra que $\beta \simeq 0.38$. Por lo tanto, la potencia de la prueba es aproximadamente $1 - \beta = 1 - 0.38 = 0.62$, que es menor que el 0.90 requerido, y se concluye en consecuencia que $n = 4$ réplicas no son suficientes. Procediendo de manera similar, puede construirse la siguiente tabla:

n	Φ^2	Φ	$a(n-1)$	β	Potencia = $(1 - \beta)$
4	4	2.00	15	0.38	0.62
5	5	2.24	20	0.18	0.82
6	6	2.45	25	0.06	0.94

Por tanto, deben correrse al menos $n = 6$ réplicas a fin de obtener una prueba con la potencia requerida.

12-6.2 Caso con efectos aleatorios

La potencia de la prueba para el modelo con efectos aleatorios es

$$\begin{aligned} 1 - \beta &= P\{\text{Se rechaza } H_0 \mid H_0 \text{ es falsa}\} \\ &= P\{F_0 > f_{\alpha,\, a-1,\, a(n-1)} \mid \sigma_\tau^2 > 0\} \end{aligned} \tag{12-34}$$

Puede demostrarse que si H_1 es verdadera ($\sigma_\tau^2 > 0$), la potencia puede calcularse usando la distribución F central, con $a - 1$ y $a(n - 1)$ grados de libertad. De hecho, el cociente

$$\frac{MS_{\text{Tratamientos}} / (\sigma^2 + n\sigma_\tau^2)}{MS_E / \sigma^2}$$

tiene la distribución F con $a - 1$ y $a(n - 1)$ grados de libertad. Entonces, por la ecuación 12-34,

$$
\begin{aligned}
1-\beta &= P\left\{\frac{MS_{\text{Tratamientos}}}{MS_E} > f_{\alpha,\,a-1,\,a(n-1)} \middle| \sigma_\tau^2 > 0\right\} \\
&= P\left\{\frac{MS_{\text{Tratamientos}}}{MS_E\left(1+n\sigma_\tau^2/\sigma^2\right)} > \frac{f_{\alpha,\,a-1,\,a(n-1)}}{\left(1+n\sigma_\tau^2/\sigma^2\right)}\right\} \\
&= P\left\{F_{a-1,\,a(n-1)} > \frac{f_{\alpha,\,a-1,\,a(n-1)}}{\left(1+n\sigma_\tau^2/\sigma^2\right)}\right\} \qquad (12\text{-}35)
\end{aligned}
$$

Este enunciado de probabilidad puede evaluarse fácilmente utilizando ciertas calculadoras de mano, o puede evaluarse usando tablas de la distribución F.

EJEMPLO 12-8

Considérese un diseño completamente aleatorizado con cinco tratamientos seleccionados al azar y seis observaciones por tratamiento. Si $\alpha = 0.05$, ¿cuál es la potencia de la prueba si $\sigma_\tau^2 = \sigma^2$?

Por la ecuación 12-35, se tiene que la potencia es

$$1-\beta = P\left\{F_{4,25} > \frac{f_{0.05,\,4,\,25}}{[1+6(1)]}\right\}$$

ya que si $\sigma_\tau^2 = \sigma^2$, el cociente $\sigma_\tau^2/\sigma^2 = 1$. Ahora bien, $f_{0.05,\,4,\,25} = 2.76$, de donde

$$
\begin{aligned}
1-\beta &= F\left\{F_{4,\,25} > \frac{2.76}{[1+6(1)]}\right\} = P\left\{F_{4,\,25} > \frac{2.76}{7}\right\} \\
&= P\{F_{4,\,25} > 0.39\} = 0.81
\end{aligned}
$$

Esta probabilidad se evaluó utilizando una calculadora que da las probabilidades de la distribución F. Puesto que la potencia de la prueba es 0.81, esto implica que la hipótesis nula H_0: $\sigma_\tau^2 = 0$ será rechazada con una probabilidad de 0.81 en esta situación experimental.

También es posible evaluar la potencia de la prueba usando las curvas de operación característica de la carta VIII del apéndice. En estas curvas se grafica la probabilidad del error tipo II, β, contra λ, donde

$$\lambda = \sqrt{1+\frac{n\sigma_\tau^2}{\sigma^2}} \qquad (12\text{-}36)$$

En el diseño de bloques aleatorizados, n se sustituye con b, el número de bloques. Puesto que por lo general no se conoce σ^2, puede usarse una estimación previa o bien definir el valor de σ_τ^2 que quiere detectarse en términos del cociente σ_τ^2/σ^2.

EJEMPLO 12-9

Considérese la situación descrita en el ejemplo 12-8. Puesto que $\sigma = 0.05$, $a = 5$, $n = 6$ y $\sigma_\tau^2 = \sigma^2$, puede encontrarse λ con la ecuación 12-36 como

$$\lambda = \sqrt{1 + 6(1)} = 2.646$$

Por la curva de operación característica con $a - 1 = 4$, $a(n - 1) = 25$ grados de libertad y $\sigma = 0.05$, se encuentra que

$$\beta \simeq 0.20$$

Por lo tanto, la potencia es aproximadamente 0.80. Esto concuerda con los resultados obtenidos en el ejemplo 12-8.

EJERCICIOS DE LA SECCIÓN 12-6

12-41. Suponga que cuatro poblaciones normales tienen la varianza común $\sigma^2 = 25$ y medias $\mu_1 = 50$, $\mu_2 = 60$, $\mu_3 = 50$ y $\mu_4 = 60$. ¿Cuántas observaciones deberán hacerse en cada población para que la probabilidad de rechazar la hipótesis de la igualdad de las medias sea de al menos 0.90? Use $\alpha = 0.05$.

12-42. Suponga que cinco poblaciones normales tienen la varianza común $\sigma^2 = 100$ y medias $\mu_1 = 175$, $\mu_2 = 190$, $\mu_3 = 160$, $\mu_4 = 200$ y $\mu_5 = 215$. ¿Cuántas observaciones por población deberán hacerse para que la probabilidad de rechazar la hipótesis de la igualdad de las medias sea de al menos 0.95? Use $\alpha = 0.01$.

12-43. Considere un experimento con un solo factor y efectos aleatorios. Si hay cuatro tratamientos y cinco observaciones por tratamiento y $\sigma_\tau^2 = 1.25\sigma^2$, encuentre la potencia de la prueba si $\alpha = 0.01$.

12-44. Considere un experimento con un solo factor y efectos aleatorios con tres tratamientos y seis observaciones por tratamiento. Encuentre la potencia de la prueba si $\sigma_\tau^2 = (\frac{3}{4})\sigma^2$ y $\alpha = 0.05$.

12-45. Considere un experimento con un solo factor y efectos aleatorios con tres tratamientos. Si $\alpha = 0.05$ y si $\sigma_\tau^2 = \sigma^2$, ¿cuántas réplicas deberán correrse si se quiere que la potencia de la prueba sea de al menos 60%?

Ejercicios complementarios

12-46. En un artículo de *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* (Vol. 15, No. 2, pp. 146-153) se describe un experimento en el que se estudió la resistencia de contacto de un relevador sólo de frenado para tres materiales diferentes (todos fueron aleaciones basadas en plata). Los datos son los siguientes:

Aleación	Resistencia de contacto									
1	95	97	99	98	99	99	99	94	95	98
2	104	102	102	105	99	102	111	103	100	103
3	119	130	132	136	141	172	145	150	144	135

a) ¿El tipo de aleación afecta la resistencia de contacto media? Use $\alpha = 0.01$.

b) Use el método gráfico de la sección 12-3.1 para determinar qué medias difieren.

c) Encuentre un intervalo de confianza de 99% para la resistencia de contacto media de la aleación 3.

d) Analice los residuales de este experimento.

12-47. En un artículo de *Lubrication Engineering* se describen los resultados de un experimento diseñado para investigar los efectos de las propiedades del material de carbono sobre la formación de burbujas en sellos con superficie de carbono. Los sellos con superficie de carbono son de uso generalizado en dispositivos como arrancadores de turbinas de aire. Se probaron cinco materiales de carbono diferentes, y se midió la irregularidad superficial. Los datos son los siguientes:

Tipo de material de carbono	Irregularidad superficial					
EC10	0.50	0.55	0.55	0.36		
EC10A	0.31	0.07	0.25	0.18	0.56	0.20
EC4	0.20	0.28	0.12			
EC1	0.10	0.16				

a) ¿El tipo de material de carbono tiene algún efecto sobre la irregularidad superficial media? Use $\alpha = 0.05$.

b) Encuentre los residuales de este experimento. ¿Una gráfica de probabilidad normal de los residuales indica algún problema con el supuesto de normalidad?

c) Grafique los residuales contra $\hat{y}_{ij}$. Comente la gráfica.

d) Encuentre un intervalo de confianza de 95% para la diferencia entre la irregularidad superficial media entre los grados de carbono EC10 Y EC1.

12-48. Aplique el método de la LSD de Fisher al experimento del ejercicio 12-47. Resumir las conclusiones referentes al efecto del tipo de material sobre la irregularidad superficial.

12-49. En un artículo de *Journal of Quality Technology* (Vol. 14, No. 2, pp. 80-89) se describe un experimento en el que se evalúan tres métodos diferentes para preparar pescado con base en criterios sensorios, asignándose una calificación de calidad. Suponga que estos métodos se han seleccionado al azar de una población grande de métodos de preparación. Los datos se muestran en la tabla del final de la página.

a) ¿Hay alguna diferencia en los métodos de preparación? Use $\alpha = 0.05$.

b) Calcule el valor P para el estadístico F del inciso a).

c) Analice los residuales de este experimento y comente la adecuación del modelo.

d) Estime los componentes de la varianza.

12-50. En un artículo de *Journal of Agricultural Engineering Research* (Vol. 52, pp. 53-76) se describe un experimento para investigar el efecto de la temperatura de secado del grano de trigo sobre la calidad del horneado del pan. Se usaron tres niveles de temperatura, y la variable de respuesta medida fue el volumen de hogazas de pan producidas. Los datos son los siguientes:

Temperatura (°C)	Volumen (CC)				
70.0	1245	1235	1285	1245	1235
75.0	1235	1240	1200	1220	1210
80.0	1225	1200	1170	1155	1095

Método	Calificación							
1	24.4	23.2	25.0	19.7	22.2	24.4	23.8	18.0
2	22.1	19.5	17.3	19.7	22.3	23.2	21.4	22.6
3	23.3	22.8	22.4	23.7	20.4	23.5	20.8	24.1

a) ¿La temperatura de secado afecta el volumen promedio de pan? Use $\alpha = 0.01$.
b) Encuentre el valor P para esta prueba.
c) Use el método gráfico de la sección 12-3.1 para determinar qué medias son diferentes.
d) Analice los residuales de este experimento y comente la adecuación del modelo.

12-51. En un artículo de *Agricultural Engineering* (pp. 672-673) se describe un experimento en el que se evalúa el aumento de peso diario de cerdos con niveles diferentes de temperatura de la porqueriza. El peso promedio de cada grupo de cerdos al principio del experimento se considera un factor perturbador. Los datos de este experimento son los siguientes:

Peso promedio (lb)	Temperatura ambiente de la porqueriza (°F) 50	60	70	80	90	100
100	1.37	1.58	2.00	1.97	1.40	0.39
150	1.47	1.75	2.16	1.82	1.14	–0.19
200	1.19	1.91	2.22	1.67	0.88	–0.77

a) ¿La temperatura ambiente de la porqueriza afecta el aumento de peso promedio? Use $\alpha = 0.05$.
b) Use el método gráfico de la sección 12-3.1 para determinar qué niveles de temperatura son diferentes.
c) Analice los residuales de este experimento y comente la adecuación del modelo.

12-52. En un artículo de *Communications of the ACM* (Vol. 30, No. 5) se estudiaron diferentes algoritmos para estimar el costo del desarrollo de software. Se aplicaron seis algoritmos a ocho proyectos de desarrollo de software y se observó el error porcentual al estimar el costo del desarrollo. Los datos se presentan en la tabla del final de la página.
a) ¿Los algoritmos difieren en la precisión de la estimación del costo promedio? Use $\alpha = 0.05$.
b) Analice los residuales de este experimento.
c) ¿Qué algoritmo se recomendaría para usarlo en la práctica?

12-53. Considere un análisis de varianza con $a = 4$ medias, $\mu_1 = 1$, $\mu_2 = 5$, $\mu_3 = 8$ y $\mu_4 = 4$. Suponga que $\sigma^2 = 4$, $n = 4$ y $\alpha = 0.05$.
a) Encuentre la potencia de la prueba F del análisis de varianza.
b) ¿De qué tamaño tendría que ser la muestra si se quisiera que la potencia de la prueba F para detectar esta diferencia de las medias fuera de al menos 0.90?

12-54. Considere un análisis de efectos aleatorios del modelo de varianza con $a = 5$ tratamientos. Sea $\sigma_\tau^2 = 2\sigma^2$, $\alpha = 0.05$, y suponga que $n = 4$.
a) Encuentre la potencia de la prueba F del análisis de varianza.
b) ¿Qué tamaño de la muestra se necesitaría si se quisiera que la potencia de la prueba F en esta situación fuera de al menos 0.90?

Algoritmo	Proyecto 1	2	3	4	5	6	7	8
1(SLIM)	1244	21	82	2221	905	839	527	122
2(COCOMO-A)	281	129	396	1306	336	910	473	199
3(COCOMO-R)	220	84	458	543	300	794	488	142
4(COCOMO-C)	225	83	425	552	291	826	509	153
5(PUNTOS DE FUNCIÓN)	19	11	–34	121	15	103	87	–17
6(ESTIMADOS)	–20	35	–53	170	104	199	142	41

EJERCICIOS PARA DESARROLLAR EL INTELECTO

12-55. Demuestre que $E(MS_E) = \sigma^2$ en el análisis de varianza del modelo de efectos fijos. ¿Cómo cambiaría la demostración si se hubiera especificado el modelo de efectos aleatorios?

12-56. Considere la prueba de la igualdad de las medias de dos poblaciones normales en las que las varianzas son desconocidas pero se suponen iguales. El procedimiento de prueba apropiado es la prueba t para dos muestras. Demuestre que la prueba t para dos muestras es equivalente a la prueba F del análisis de varianza con un solo factor.

12-57. Considere el análisis de varianza con $a = 2$ tratamientos. Demuestre que MS_E de este análisis es igual a la estimación de la varianza agrupada usada en la prueba t para dos muestras.

12-58. Demuestre que la varianza de la combinación lineal

$$\sum_{i=1}^{a} c_i Y_{i\cdot} \text{ es } \sum_{i=1}^{a} n_i c_i^2.$$

12-59. Suponga que en un modelo de efectos fijos hay n observaciones para cada uno de cuatro tratamientos. Sean Q_1^2, Q_2^2 y Q_3^2 las sumas de cuadrados con un grado de libertad para los contrastes ortogonales. Demuestre que $SS_{\text{Tratamientos}} = Q_1^2 + Q_2^2 + Q_3^2$.

12-60. Considere el diseño completamente aleatorizado con un solo factor, con a tratamientos y n réplicas. Demuestre que si la diferencia entre las medias de dos tratamientos cualesquiera es tan grande como D, entonces el valor mínimo que puede asumir el parámetro Φ^2 de la curva OC es

$$\Phi^2 = \frac{nD^2}{2a\sigma^2}$$

12-61. Considere el diseño completamente aleatorizado con un solo factor. Demuestre que un intervalo de confianza del $100(1-\alpha)$ por ciento para σ^2 es

$$\frac{(N-a)MS_E}{\chi^2_{\alpha/2,\,N-a}} \le \sigma^2 \le \frac{(N-a)MS_E}{\chi^2_{1-\alpha/2,\,N-a}}$$

donde N es el número total de observaciones del diseño experimental.

12-62. Considere el modelo de efectos aleatorios para el diseño completamente aleatorizado con un solo factor. Demuestre que un intervalo de confianza del $100(1-\alpha)$ por ciento para el cociente de los componentes de la varianza σ_τ^2/σ^2 está dado por

$$L \le \frac{\sigma_\tau^2}{\sigma^2} \le U$$

donde

$$L = \frac{1}{n}\left[\frac{MS_{\text{Tratamientos}}}{MS_E} \times \left(\frac{1}{f_{\alpha/2,\,a-1,\,N-a}}\right) - 1\right]$$

y

$$U = \frac{1}{n}\left[\frac{MS_{\text{Tratamientos}}}{MS_E} \times \left(\frac{1}{f_{1-\alpha/2,\,a-1,\,N-a}}\right) - 1\right]$$

12-63. Considere un modelo de efectos aleatorios para el diseño completamente aleatorizado con un solo factor. Demuestre que un intervalo de confianza del $100(1-\alpha)$ por ciento para el cociente de los componentes de la varianza $\sigma_\tau^2/(\sigma^2+\sigma_\tau^2)$ es

$$\frac{L}{1+L} \le \frac{\sigma_\tau^2}{\sigma^2+\sigma_\tau^2} \le \frac{U}{1+U}$$

EJERCICIOS PARA DESARROLLAR EL INTELECTO

donde L y U se definen como en el ejercicio 12-62.

12-64. Continuación del ejercicio 12-63. Use los resultados del ejercicio 12-63 anterior para encontrar una intervalo de confianza del $100(1 - \alpha)$ por ciento para $\sigma^2/(\sigma^2 + \sigma_\tau^2)$.

12-65. Considere el modelo de efectos fijos del diseño completamente aleatorizado con un solo factor. Los parámetros del modelo están sujetos a la restricción $\sum_{i=1}^{a} \tau_i = 0$. (De hecho, pudieron usarse otras restricciones, pero ésta es simple y da como resultado estimaciones intuitivamente accesibles de los parámetros del modelo.) Para el caso de tamaños de las muestras $n_1, n_2, \ldots, n_a$ diferentes, la restricción es $\sum_{i=1}^{a} n_i \tau_i = 0$. Use esto para demostrar que

$$E(MS_{\text{Tratamientos}}) = \sigma^2 + \frac{\sum_{i=1}^{a} n_i \tau_i^2}{a-1}$$

¿Esto sugiere que la hipótesis nula de este modelo es H_0: $n_1\tau_1 = n_2\tau_2 = \ldots = n_a\tau_a = 0$?

12-66. Determinación del tamaño de la muestra. En el diseño completamente aleatorizado con un solo factor, la precisión de un intervalo de confianza del $100(1 - \alpha)$ por ciento para la diferencia de las medias de dos tratamientos cualesquiera es $t_{\alpha/2,\, a(n-1)}\sqrt{2MS_E/n}$.

a) Demuestre que si A es la precisión deseada del intervalo, entonces el tamaño de la muestra requerido es

$$n = \frac{2F_{\alpha/2,\, 1,\, a(n-1)} MS_E}{A^2}$$

b) Suponga que al comparar $a = 5$ medias se tiene una estimación preliminar de σ^2 de 4. Si se quiere que el intervalo de confianza de 95% para la diferencia de las medias tenga una precisión de 2, ¿cuántas réplicas deberán hacerse?

13 Diseño de experimentos con varios factores

PLAN GENERAL DEL CAPÍTULO

13-1 INTRODUCCIÓN

13-2 ALGUNAS APLICACIONES DE LAS TÉCNICAS DE DISEÑO EXPERIMENTAL

13-3 EXPERIMENTOS FACTORIALES

13-4 EXPERIMENTOS FACTORIALES CON DOS FACTORES

- 13-4.1 Análisis estadístico del modelo con efectos fijos
- 13-4.2 Verificación de la adecuación del modelo
- 13-4.3 Salida de computadora
- 13-4.4 Una observación por celda
- 13-4.5 Factores aleatorios

13-5 EXPERIMENTOS FACTORIALES GENERALES

13-6 DISEÑO FACTORIAL 2^k

- 13-6.1 Diseño 2^2
- 13-6.2 Diseño 2^k para $k \geq 3$ factores
- 13-6.3 Una sola réplica del diseño 2^k
- 13-6.4 Adición de los puntos centrales a un diseño 2^k

13-7 FORMACIÓN DE BLOQUES Y CONFUSIÓN EN EL DISEÑO 2^k

13-8 RÉPLICA FRACCIONADA DEL DISEÑO 2^k

- 13-8.1 Media fracción del diseño 2^k
- 13-8.2 Fracciones menores: el factorial fraccionado 2^{k-p}

13-9 MÉTODOS Y DISEÑOS DE SUPERFICIES DE RESPUESTA

- 13-9.1 Método del ascenso más pronunciado
- 13-9.2 Análisis de una superficie de respuesta de segundo orden

13-1 INTRODUCCIÓN

Un experimento no es sino una prueba o una serie de pruebas. En todas las áreas de la ingeniería y las ciencias se llevan a cabo experimentos, y constituyen un recurso primordial para conocer la forma en que operan los sistemas y los procesos. La validez de las conclusiones que se sacan de un experimento dependen en gran medida de la manera en que éste se llevó a cabo. Por lo tanto, el **diseño** del experimento desempeña un papel fundamental en la eventual solución del problema que dio lugar al experimento.

El presente capítulo se enfoca en los experimentos que incluyen dos o más factores que el experimentador considera de posible importancia. Se introducirá el **diseño experimental factorial** como una técnica poderosa para este tipo de problema. En general, en un diseño experimental factorial los ensayos (o corridas) experimentales se realizan para todas las combinaciones de los niveles de los factores. Por ejemplo, si un ingeniero químico tiene interés en investigar los efectos del tiempo de reacción y la temperatura de reacción sobre el rendimiento de un proceso, y si se consideran importantes dos niveles de tiempo (1 hr y 1.5 hr) y dos niveles de temperatura (125°F y 150°F), entonces un experimento factorial consistiría en realizar las corridas experimentales para cada una de las cuatro combinaciones posibles de estos niveles del tiempo de reacción y la temperatura de reacción.

La mayoría de los conceptos estadísticos que se introdujeron en el capítulo 12 para experimentos con un solo factor pueden extenderse a los experimentos factoriales de este capítulo. En particular, se seguirá usando el análisis de varianza como una de las principales herramientas estadísticas para el análisis de datos estadísticos. Se introducirán además varios métodos gráficos que son útiles en el análisis de datos de experimentos diseñados.

13-2 ALGUNAS APLICACIONES DE LAS TÉCNICAS DE DISEÑO EXPERIMENTAL

El diseño experimental es una herramienta de gran utilidad para ingenieros y científicos que tienen interés en mejorar el desempeño de un proceso de manufactura. También tiene una aplicación generalizada en el desarrollo de nuevos procesos y en el diseño de nuevos productos. Se presentan a continuación algunos ejemplos.

EJEMPLO 13-1

Experimento de caracterización de un proceso

Un equipo de ingenieros de desarrollo trabaja en un nuevo proceso para soldar componentes electrónicos en tarjetas de circuitos impresos. Específicamente, el equipo trabaja con un nuevo tipo de máquina de soldadura líquida que deberá reducir el número de juntas de soldadura defectuosas. (Una máquina de soldadura líquida precalienta las tarjetas de circuitos impresos y después las pone en contacto con una onda de soldadura líquida. Esta máquina hace todas las conexiones eléctricas y la mayor parte de las conexiones mecánicas de los componentes en la tarjeta del circuito impreso. Los defectos de soldadura requieren correcciones o repetir la operación, lo cual aumenta el costo y con frecuencia daña las tarjetas.) El proceso tendrá varias (quizá muchas) variables, las cuales por lo general no tienen la misma importancia. La lista inicial de las variables candidatas que deberán incluirse en el experimento se hace combinando

el conocimiento y la información acerca del proceso de todos los integrantes del equipo. En este ejemplo, los ingenieros efectuaron una sesión de *lluvia de ideas* e invitaron a participar al personal de manufactura con experiencia en el uso de varios tipos de equipo de soldadura líquida. El equipo determinó que la máquina de soldadura líquida tiene varias variables que pueden controlarse. Éstas son:

1. Temperatura de la soldadura
2. Temperatura del precalentamiento
3. Velocidad de la transportadora
4. Tipo de flujo
5. Peso específico del flujo
6. Profundidad de la onda de soldadura
7. Ángulo de la transportadora

Además de estos factores controlables, hay varios factores que no es sencillo controlar una vez que la máquina entra en operación, entre los que se encuentran:

1. Espesor de la tarjeta del circuito impreso
2. Tipos de componentes usados en la tarjeta
3. Disposición de los componentes en la tarjeta
4. Operador
5. Factores ambientales
6. Volumen de producción

En ocasiones, a los factores que no pueden controlarse se les llama factores de *ruido*. En la figura 13-1 se muestra una representación esquemática del proceso.

En esta situación, al ingeniero le interesa **caracterizar** la máquina de soldadura líquida; es decir, le interesa determinar qué factores (tanto los controlables como los no controlables) afectan la ocurrencia de defectos en las tarjetas de circuitos impresos. Para determinar estos factores, el ingeniero puede diseñar un experimento que le permita estimar la magnitud y la dirección de los efectos de los factores. En ocasiones a un experimento como éste se le llama **experimento tamiz, de selección** o **exhaustivo**. La información de este estudio de caracterización o experimento tamiz puede ayudar a determinar las variables críticas del proceso, así como la dirección del ajuste de estos factores a fin de reducir el número de defectos, y contribuir a determinar qué variables del proceso deberán someterse a control estricto durante la manufactura para evitar que ocurran niveles elevados de defectos y un desempeño errático del proceso.

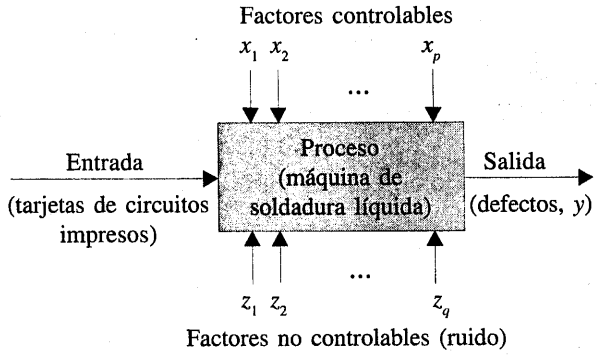


Figura 13-1 El experimento de la máquina de soldadura líquida.

EJEMPLO 13-2

Un experimento de optimización

En un experimento de caracterización, el interés se encuentra en determinar qué factores afectan la respuesta. El siguiente paso lógico es determinar la región de los factores importantes que produce la **respuesta óptima**. Por ejemplo, si la respuesta es el costo, se buscará una región de costo mínimo.

Como ilustración, supóngase que en el rendimiento de un proceso químico influyen la temperatura de operación y el tiempo de reacción. El proceso se opera actualmente a 155°F y 1.7 horas de tiempo de reacción, y el rendimiento actual del proceso es de cerca de 75%. En la figura 13-2 se muestra una vista desde arriba del espacio tiempo-temperatura. En esta gráfica los puntos que tienen un rendimiento constante se han unido con líneas. Estas líneas son los **contornos** del rendimiento, y se ilustran los contornos con un rendimiento de 60, 70, 80, 90 y 95%. Para localizar el óptimo, podría empezarse con un experimento factorial como el que se describió en el ejemplo 13-1, aplicando los dos factores, el tiempo y la temperatura, a 10°F y 0.5 horas arriba y abajo de las condiciones de operación actuales. Este diseño factorial con dos factores se muestra en la figura 13-2. Las respuestas promedio observadas en los cuatro puntos del experimento (145°F, 1.2 hr; 145°F, 2.2 hr; 165°F, 1.2 hr, y 165°F, 2.2 hr) indican que debería procederse en la dirección general del aumento de la temperatura y la disminución del tiempo de reacción a fin de incrementar el rendimiento. Podrían hacerse experimentos adicionales en esta dirección para localizar la región del rendimiento máximo.

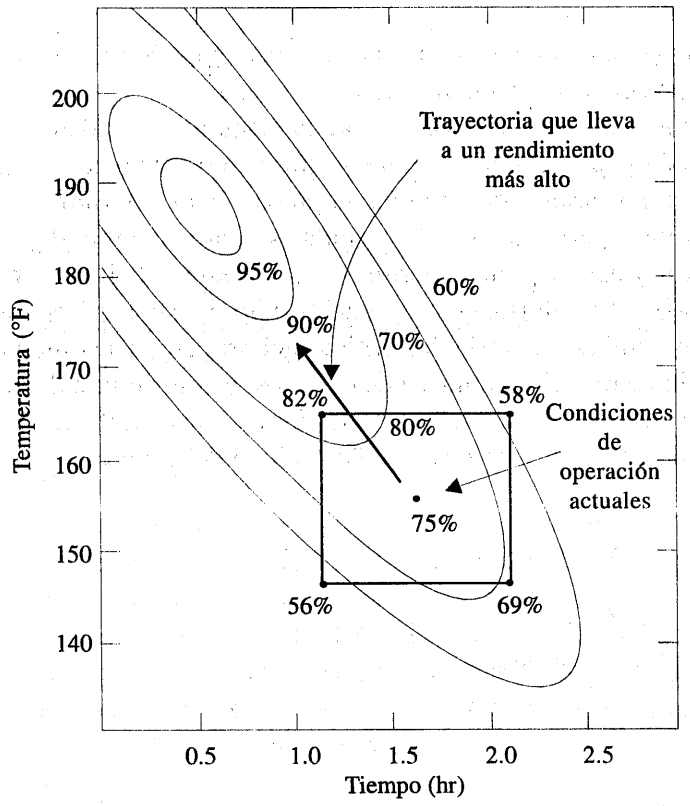


Figura 13-2 Gráfica de contorno del rendimiento como una función del tiempo de reacción y la temperatura de reacción, la cual ilustra un experimento de optimización.

EJEMPLO 13-3

Ejemplo del diseño de un producto

El diseño experimental también puede usarse para desarrollar productos nuevos. Por ejemplo, supóngase que un grupo de ingenieros está diseñando el gozne de la puerta de un automóvil. La característica del producto es el esfuerzo amortiguador, o la capacidad de retención del tope que impide que la puerta se cierre cuando el vehículo se estaciona en una pendiente. El mecanismo amortiguador consta de un resorte de hojas y un rodillo. Cuando la puerta se abre, el rodillo se desplaza por un arco que hace que el resorte de hojas se comprima. Para cerrar la puerta es necesario vencer la fuerza del resorte, la cual produce el esfuerzo amortiguador. El equipo de ingenieros piensa que el esfuerzo amortiguador es una función de los siguientes factores:

1. La distancia que se desplaza el cilindro
2. La altura del resorte del pivote a la base
3. La distancia horizontal del pivote al resorte
4. La altura libre del resorte auxiliar
5. Altura libre del resorte principal

Los ingenieros pueden construir un prototipo del mecanismo del gozne en el que es posible variar todos estos factores dentro de ciertos rangos. Una vez que se han identificado los niveles apropiados de estos cinco factores, puede diseñarse un experimento compuesto por varias combinaciones de los niveles de los factores, y el prototipo puede probarse con estas combinaciones. Esto producirá información referente a los factores que tienen una influencia mayor sobre el esfuerzo amortiguador del tope y, mediante el análisis de esta información, puede mejorarse el diseño del tope.

En los ejemplos anteriores sólo se ilustran tres aplicaciones de los métodos del diseño experimental. En el campo de la ingeniería son numerosas las aplicaciones del diseño experimental. Algunas áreas de uso potencial son:

1. Corrección del proceso
2. Desarrollo y optimización del proceso
3. Evaluación de materiales y alternativas
4. Confiabilidad y pruebas de duración
5. Pruebas de desempeño
6. Configuración del diseño de productos
7. Determinación de la tolerancia de componentes

Los métodos de diseño experimental permiten resolver con eficiencia estos problemas en las fases iniciales del ciclo del producto. Esto ofrece el potencial de una reducción radical del costo del producto y del tiempo de desarrollo para introducirlo en el mercado.

13-3 EXPERIMENTOS FACTORIALES

Cuando en un experimento hay varios factores de interés, deberá usarse un **diseño experimental factorial**. Como ya se señaló, en estos experimentos los factores se hacen variar en conjunto.

Definición

Por **experimento factorial** se entiende que en cada ensayo completo o réplica del experimento se investigan todas las combinaciones posibles de los niveles de los factores.

Por tanto, si hay dos factores A y B con a niveles del factor A y b niveles del factor B, entonces cada réplica incluye todas las combinaciones de tratamientos ab.

El efecto de un factor se define como el cambio en la respuesta que produce un cambio en el nivel del factor. A éste se le llama **efecto principal** porque se refiere a los factores primarios del estudio. Por ejemplo, considérense los datos de la tabla 13-1. Se trata de un experimento factorial con dos factores, A y B, cada uno con dos niveles (A_{bajo}, A_{alto}, B_{bajo}, B_{alto}). El efecto principal del factor A es la diferencia entre la respuesta promedio en el nivel alto de A y la respuesta promedio en el nivel bajo de A, o

$$A = \frac{30+40}{2} - \frac{10+20}{2} = 20$$

Es decir, cambiar el factor A del nivel bajo al nivel alto produce un incremento de la respuesta media de 20 unidades. De manera similar, el efecto principal de B es

$$B = \frac{20+40}{2} - \frac{10+30}{2} = 10$$

En algunos experimentos, la diferencia de la respuesta entre los niveles de un factor no es la misma en todos los niveles de los otros factores. Cuando esto ocurre, existe una **interacción** entre los factores. Por ejemplo, considérense los datos de la tabla 13-2. En el nivel bajo del factor B, el efecto de A es

$$A = 30 - 10 = 20$$

y en el nivel alto del factor B, el efecto de A es

$$A = 0 - 20 = -20$$

Tabla 13-1 Experimento factorial con dos factores

Factor A	Factor B	
	B_{bajo}	B_{alto}
A_{baja}	10	20
A_{alta}	30	40

Tabla 13-2 Experimento factorial con interacción

Factor A	Factor B	
	B_{bajo}	B_{alto}
A_{bajo}	10	20
A_{alto}	30	0

Puesto que el efecto de A depende del nivel elegido del factor B, hay interacción entre A y B.

Cuando la interacción es grande, los efectos principales correspondientes tienen muy poco significado práctico. Por ejemplo, utilizando los datos de la tabla 13-2, se encuentra que el efecto principal de A es

$$A = \frac{30+0}{2} - \frac{10+20}{2} = 0$$

y uno estaría tentado a concluir que el factor A no tiene efecto alguno. Sin embargo, cuando se examinaron los efectos de A en *niveles diferentes del factor B*, se vio que no era éste el caso. El efecto del factor A depende de los niveles del factor B. Por tanto, el conocimiento de la interacción AB es más útil que el conocimiento del efecto principal. Una interacción significativa puede enmascarar la significación de los efectos principales. Por consiguiente, cuando está presente una interacción, los efectos principales de los factores que intervienen en la interacción quizá no tengan mucho significado.

Es sencillo estimar el efecto de una interacción en experimentos factoriales como los que se ilustraron en las tablas 13-1 y 13-2. En este tipo de experimentos, cuando ambos factores tienen dos niveles, el efecto de la interacción AB es la diferencia de los promedios de las diagonales. Esto representa la mitad de la diferencia entre los efectos de A en los dos niveles de B. Por ejemplo, en la tabla 13-1 se encuentra que el efecto de la interacción AB es

$$AB = \frac{20+30}{2} - \frac{10+40}{2} = 0$$

Por tanto, no hay interacción entre A y B. En la tabla 13-2, el efecto de la interacción AB es

$$AB = \frac{20+30}{2} - \frac{10+0}{2} = 20$$

Como ya se señaló, el efecto de la interacción en estos datos es muy grande.

Hay varias maneras de ilustrar gráficamente el concepto de interacción. En la figura 13-3 se grafican los datos de la tabla 13-1 contra los niveles de A para los dos niveles de B. Obsérvese que las rectas B_{baja} y B_{alta} son aproximadamente paralelas, indicando que los factores A y B no interactúan de manera significativa. En la figura 13-4 se presenta una gráfica similar para los

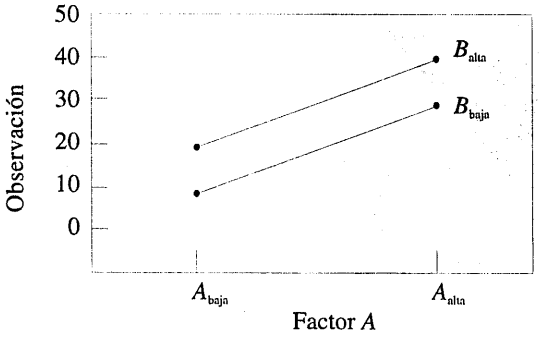


Figura 13-3 Experimento factorial, sin interacción.

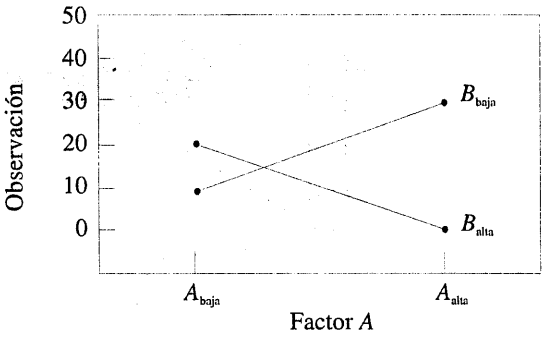


Figura 13-4 Experimento factorial, con interacción.

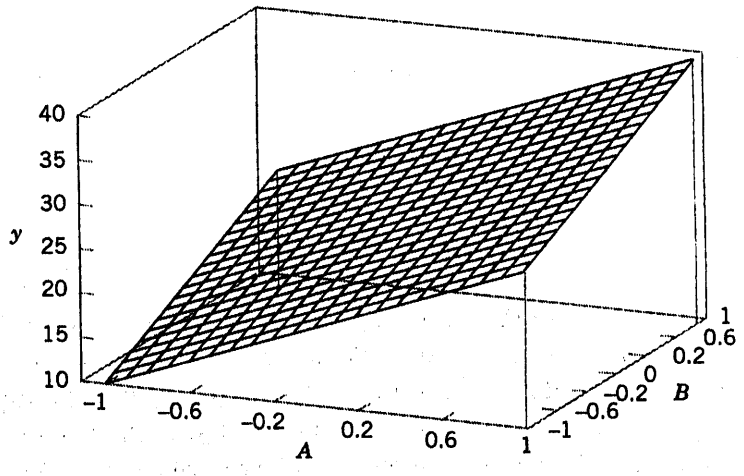


Figura 13-5 Gráfica de superficie tridimensional de los datos de la tabla 13-1, donde se muestran los efectos principales de los dos factores A y B.

datos de la tabla 13-2. En esta gráfica, las rectas B_{baja} y B_{alta} no son paralelas, indicando la interacción entre los factores A y B. A estas representaciones gráficas se les llama **gráficas de interacción de dos factores**. Con frecuencia son útiles para presentar los resultados de experimentos, y muchos programas de software de computadora que se usan para analizar datos de experimentos diseñados construirán estas gráficas automáticamente.

En las figura 13-5 y 13-6 se presenta otra ilustración gráfica de los datos de las tablas 13-1 y 13-2. En la figura 13-5 se muestra una **gráfica de superficie tridimensional** de los datos de la tabla 13-1. Estos datos no contienen ninguna interacción, y la gráfica de superficie es un plano que se localiza arriba del espacio A-B. La pendiente del plano en las direcciones A y B es proporcional a los efectos principales de los factores A y B, respectivamente. La figura 13-6 es una gráfica de superficie de los datos de la tabla 13-2. Obsérvese que el efecto de la interacción en

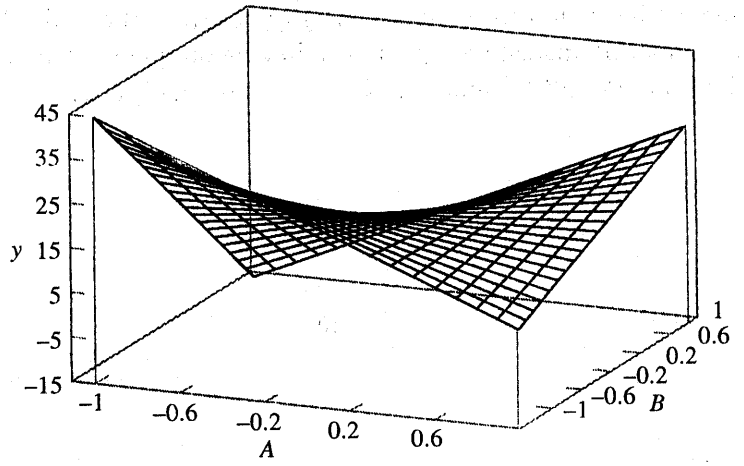


Figura 13-6 Gráfica de superficie tridimensional de los datos de la tabla 13-2, donde se muestra el efecto de la interacción de A y B.

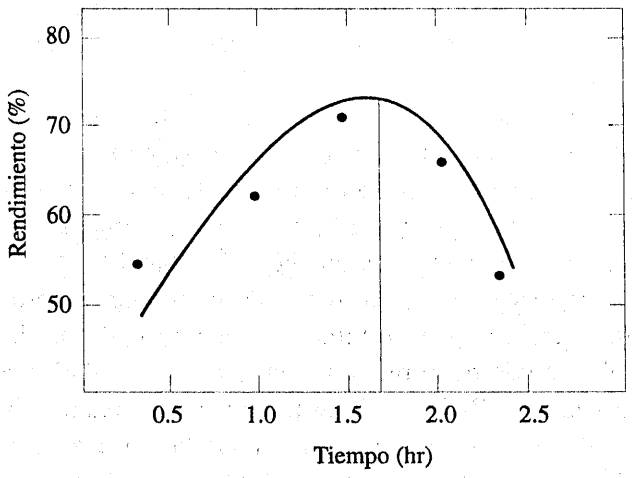


Figura 13-7 Rendimiento contra tiempo de reacción con temperatura constante de 155°F.

estos datos es "torcer" el plano, por lo que hay curvatura en la función de respuesta. **Los experimentos factoriales son la única manera de descubrir interacciones entre variables.**

Una alternativa del diseño factorial que (desafortunadamente) se usa en la práctica consiste en cambiar los factores *uno a la vez* en lugar de hacerlos variar simultáneamente. Para ilustrar este procedimiento de un factor a la vez, considérese el experimento de optimización descrito en el ejemplo 13-2. El ingeniero tiene interés en encontrar los valores de la temperatura y del tiempo que maximizan el rendimiento. Suponga que la temperatura se fija en 155°F (el nivel de operación actual) y que se realizan cinco corridas con diferentes niveles de tiempo, por ejemplo, 0.5 horas, 1.0 horas, 2.0 horas y 2.5 horas. Los resultados de estas corridas se muestran en la figura 13-7. Esta figura indica que el rendimiento máximo se alcanza en cerca de 1.7 horas de tiempo de reacción. Para optimizar la temperatura, el ingeniero fija el tiempo en 1.7 horas (el óptimo aparente) y realiza cinco corridas con diferentes temperaturas, por ejemplo, 140°F, 150°F, 160°F, 170°F y 180°F. Los resultados de este conjunto de corridas se grafican en la figura 13-8. El

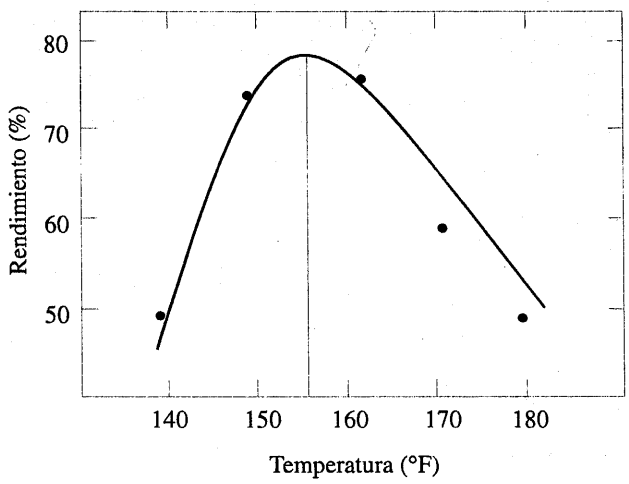


Figura 13-8 Rendimiento contra temperatura con tiempo de reacción constante de 1.7 horas.

rendimiento máximo ocurre en cerca de 155°F. Por lo tanto, se concluiría que operar el proceso en 155°F y 1.7 horas es el mejor ajuste de las condiciones de operación, dando como resultado rendimientos de cerca de 75%.

En la figura 13-9 se presenta la gráfica de contorno del rendimiento como una función de la temperatura y del tiempo con los experimentos realizados con un factor a la vez superpuestos en los contornos. Evidentemente, este enfoque de un factor a la vez ha fallado dramáticamente en este caso, ya que el verdadero óptimo está al menos 20 puntos de rendimiento más arriba con tiempos de reacción mucho más bajos y temperaturas mucho más altas. El fracaso para descubrir la importancia de los tiempos de reacción más cortos es de particular importancia, ya que esto podría tener un impacto significativo en el volumen o capacidad de producción, en la planeación de la producción, en el costo de manufactura y en la productividad total.

El enfoque de un factor a la vez ha fallado aquí porque no puede detectar la interacción entre la temperatura y el tiempo. Los experimentos factoriales son la única manera de detectar interacciones. Además, el método de un factor a la vez es ineficiente. Requerirá un número mayor de experimentos que el método factorial y, como se ha visto, no existe la seguridad de que producirá los resultados correctos.

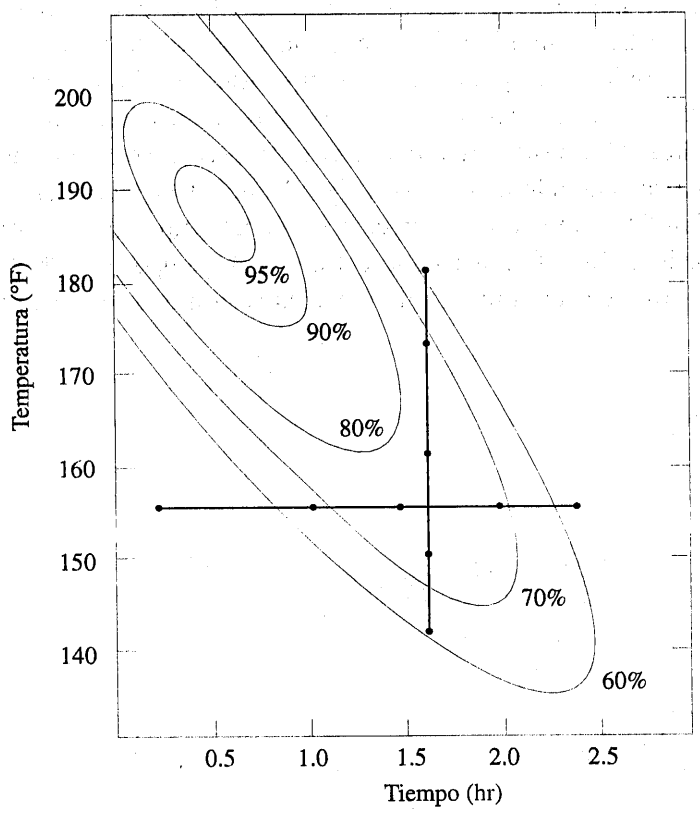


Figura 13-9 Experimento de optimización usando el método de un factor a la vez.

13-4 EXPERIMENTOS FACTORIALES CON DOS FACTORES

El experimento factorial más sencillo es en el que intervienen solamente dos factores, por ejemplo, A y B. Hay a niveles del factor A y b niveles del factor B. En la tabla 13-3 se muestra este diseño factorial de dos factores. El experimento tiene n **réplicas** y cada réplica contiene todas las combinaciones de tratamientos ab. La observación de la celda ij-ésima para la réplica k-ésima se denota por y_{ijk}. Cuando se realiza el experimento, las abn observaciones se correrían en **orden aleatorio**. Por tanto, como en el experimento con un solo factor estudiado en el capítulo 12, el factorial con dos factores es un *diseño completamente aleatorizado.*

Las observaciones pueden describirse con el modelo estadístico lineal

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{13-1}$$

donde μ es el efecto promedio global, τ_i es el efecto del nivel i-ésimo del factor A, β_j es el efecto del nivel j-ésimo del factor B, $(\tau\beta)_{ij}$ es el efecto de la interacción entre A y B, y ε_{ijk} es un componente del error aleatorio que tiene una distribución normal con media cero y varianza σ^2. El interés se encuentra en probar las hipótesis de que no hay efecto principal para el factor A, que no hay efecto principal para B, y que no hay efecto de interacción AB. Como en el caso de los experimentos con un solo factor del capítulo 12, se usará el análisis de varianza para probar estas hipótesis. Puesto que hay dos factores en el experimento, al procedimiento de prueba en ocasiones se le llama análisis de varianza de dos direcciones.

13-4.1 Análisis estadístico del modelo con efectos fijos

Suponga que A y B son **factores fijos**. Es decir, el experimentador elige específicamente los a niveles del factor A y los b niveles del factor B, y las inferencias se restringen únicamente a estos

Tabla 13-3 Disposición de los datos en un diseño factorial de dos factores

		Factor B				
		1	2	... b	Totales	Promedios
Factor A	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$	$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$	$y_{1\cdot\cdot}$	$\bar{y}_{1\cdot\cdot}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$	$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$	$y_{2\cdot\cdot}$	$\bar{y}_{2\cdot\cdot}$
	$\vdots$					
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$	$y_{ab1}, y_{ab2}, \ldots, y_{abn}$	$y_{a\cdot\cdot}$	$\bar{y}_{a\cdot\cdot}$
Promedios totales		$y_{\cdot 1\cdot}$ $\bar{y}_{\cdot 1\cdot}$	$y_{\cdot 2\cdot}$ $\bar{y}_{\cdot 2\cdot}$	$y_{\cdot b\cdot}$ $\bar{y}_{\cdot b\cdot}$	$y_{\cdot\cdot\cdot}$	$\bar{y}_{\cdot\cdot\cdot}$

niveles. En este modelo, se acostumbra definir los efectos τ_i, β_j y $(\tau\beta)_{ij}$ como desviaciones de la media, por lo que $\sum_{i=1}^{a} \tau_i = 0, \sum_{j=1}^{b} \beta_j = 0, \sum_{i=1}^{a} (\tau\beta)_{ij} = 0$ y $\sum_{j=1}^{b} (\tau\beta)_{ij} = 0$.

Puede usarse el **análisis de varianza** para probar hipótesis relativas a los efectos principales de los factores A y B y la interacción AB. Para presentar los detalles del análisis de varianza, se necesitarán algunos símbolos, algunos de los cuales se ilustran en la tabla 13-3. Sea que $y_{i\cdot\cdot}$ denote el total de las observaciones tomadas en el nivel i-ésimo del factor A; que $y_{\cdot j\cdot}$ denote el total de las observaciones tomadas en el nivel j-ésimo del factor B; que $y_{ij\cdot}$ denote el total de las observaciones en la celda ij-ésima de la tabla 13-3; y que $y_{\cdots}$ denote el gran total de todas las observaciones. Se definen $\bar{y}_{i\cdot\cdot}$, $\bar{y}_{\cdot j\cdot}$, $\bar{y}_{ij\cdot}$, y $\bar{y}_{\cdots}$ como los promedios de los renglones, las columnas, las celdas y el gran promedio correspondientes. Es decir,

$$y_{i\cdot\cdot} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{i\cdot\cdot} = \frac{y_{i\cdot\cdot}}{bn} \qquad i = 1, 2, \ldots, a$$

$$y_{\cdot j\cdot} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{\cdot j\cdot} = \frac{y_{\cdot j\cdot}}{an} \qquad j = 1, 2, \ldots, b$$

$$y_{ij\cdot} = \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{ij\cdot} = \frac{y_{ij\cdot}}{n} \qquad \begin{array}{l} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{array}$$

$$y_{\cdots} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{\cdots} = \frac{y_{\cdots}}{abn}$$

Las hipótesis que se probarán son las siguientes:

1. H_0: $\tau_1 = \tau_2 = \cdots = \tau_a = 0$ (ningún efecto principal del factor A)
   H_1: al menos una $\tau_i \neq 0$
2. H_0: $\beta_1 = \beta_2 = \cdots = \beta_b = 0$ (ningún efecto principal del factor B) (13-2)
   H_1: al menos una $\beta_j \neq 0$
3. H_0: $(\tau\beta)_{11} = (\tau\beta)_{12} = \cdots = (\tau\beta)_{ab} = 0$ (ninguna interacción)
   H_1: al menos una $(\tau\beta)_{ij} \neq 0$

Como anteriormente, el análisis de varianza prueba estas hipótesis descomponiendo la variabilidad total de los datos en partes componentes y comparando después los diferentes elementos de esta descomposición. La variabilidad total se mide por la suma total de cuadrados de las observaciones

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \left(y_{ijk} - y_{\cdots}\right)^2$$

y la descomposición de la suma de cuadrados es como se define a continuación.

La identidad de la suma de cuadrados para un análisis de varianza de dos factores es

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{\cdots}\right)^2 &= bn\sum_{i=1}^{a}\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)^2\\ &+an\sum_{j=1}^{b}\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)^2\\ &+n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)^2\\ &+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij\cdot}\right)^2 \qquad (13\text{-}3)\end{aligned}$$

La demostración de la identidad de esta suma de cuadrados puede establecerse escribiendo la suma total de cuadrados como

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{\cdots}\right)^2 &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left[\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)+\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)\right.\\ &\left.+\left(y_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)+\left(y_{ijk}-\bar{y}_{ij\cdot}\right)\right]^2\\ &= bn\sum_{i=1}^{a}\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)^2+an\sum_{j=1}^{b}\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)^2\\ &+n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)^2\\ &+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij\cdot}\right)^2\\ &+\text{6 términos de productos cruzados}\end{aligned}$$

Sin embargo, todos los términos de los productos cruzados son cero, por lo que esta última ecuación se reduce a la identidad de la suma de cuadrados de la ecuación 13-3. Por tanto, se hace la partición de la suma total de cuadrados en una suma de cuadrados debida a los "renglones" o factor A (SS_A), una suma de cuadrados debida a las "columnas" o factor B (SS_B), una suma de cuadrados debida a la interacción entre A y B (SS_{AB}), y una suma de cuadrados debida al error (SS_E). Obsérvese que debe haber al menos dos réplicas para obtener una suma de cuadrados del error diferente de cero.

Utilizando símbolos, la identidad de la suma de cuadrados de la ecuación 13-3 puede escribirse como

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E \qquad (13\text{-}4)$$

Hay en total $abn - 1$ grados de libertad. Los efectos principales de A y B tienen $a - 1$ y $b - 1$ grados de libertad, mientras que el efecto de la interacción AB tiene $(a - 1)(b - 1)$ grados de libertad. En cada una de las celdas ab de la tabla 13-3 hay $n - 1$ grados de libertad entre las n réplicas, y las observaciones de la misma celda sólo pueden diferir por el error aleatorio. Por lo tanto, hay $ab\,(n - 1)$ grados de libertad del error. Por lo tanto, la partición de los grados de libertad se hace de acuerdo con

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)\,(b - 1) + ab\,(n - 1)$$

Si cada una de las sumas de cuadrados del segundo miembro de la ecuación 13-4 se divide por el número correspondiente de grados de libertad, se obtienen los cuadrados medios de A, B, la interacción y el error:

$$MS_A = \frac{SS_A}{a-1} \qquad MS_B = \frac{SS_B}{b-1}$$

$$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)} \qquad MS_E = \frac{SS_E}{ab\,(n-1)}$$

Suponiendo que los factores A y B son fijos, no es difícil demostrar que los valores esperados de estos cuadrados medios son

$$E\left(MS_A\right) = E\left(\frac{SS_A}{a-1}\right) = \sigma^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E\left(MS_B\right) = E\left(\frac{SS_B}{b-1}\right) = \sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$E\left(MS_{AB}\right) = E\left(\frac{SS_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{(a-1)(b-1)}$$

$$E\left(MS_E\right) = E\left(\frac{SS_E}{ab\,(n-1)}\right) = \sigma^2$$

Al examinar estos cuadrados medios esperados, es claro que si las hipótesis nulas sobre los efectos principales H_0: $\tau_i = 0$, H_0: $\beta_j = 0$, y la hipótesis de la interacción H_0: $(\tau\beta)_{ij} = 0$ son todas verdaderas, los cuatro cuadrados medios son estimaciones insesgadas de σ^2.

Para probar que todos los efectos de los factores de los renglones son iguales a cero (H_0: $\tau_i = 0$), se usaría el cociente

$$F_0 = \frac{MS_A}{MS_E}$$

que tiene una distribución F con $a - 1$ y $ab\,(n - 1)$ grados de libertad si H_0: $\tau_i = 0$ es verdadera. Esta hipótesis nula se rechaza con el nivel de significación α si $f_0 > f_{\alpha,\, a-1,\, ab(n-1)}$. De manera similar, para probar la hipótesis de que todos los efectos de los factores de las columnas son cero (H_0: $\beta_j = 0$), se usaría el cociente

$$F_0 = \frac{MS_B}{MS_E}$$

que tiene una distribución F con $b - 1$ y $ab\,(n - 1)$ grados de libertad si H_0: $\beta_j = 0$ es verdadera. Esta hipótesis nula se rechaza con el nivel de significación α si $f_0 > f_{\alpha,\, b-1,\, ab(n-1)}$. Por último, para probar la hipótesis H_0: $(\tau\beta)_{ij} = 0$, que es la hipótesis de que todos los efectos de interacción son cero, se usa el cociente

$$F_0 = \frac{MS_{AB}}{MS_E}$$

que tiene una distribución F con $(a - 1)(b - 1)$ y $ab\,(n - 1)$ grados de libertad si la hipótesis nula H_0: $(\tau\beta)_{ij} = 0$. Esta hipótesis se rechaza con el nivel de significación α si $f_0 > f_{\alpha,\, (a-1)(b-1),\, ab(n-1)}$.

Por lo general, es conveniente realizar primero la prueba de la interacción y evaluar después los efectos principales. Si la interacción no es significativa, la interpretación de las pruebas de los efectos principales es directa. Sin embargo, como se señaló en la sección 13-3, cuando la interacción es significativa, quizá los efectos principales de los factores que intervienen en la interacción no sean de gran valor en la interpretación práctica. El conocimiento de la interacción suele ser más importante que el conocimiento de los efectos principales.

Las fórmulas para calcular las sumas de cuadrados se obtienen con facilidad.

Definición

Fórmulas para calcular las sumas de los cuadrados en un análisis de varianza de dos factores.

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn} \tag{13-5}$$

$$SS_A = \sum_{i=1}^{a} \frac{y_{i..}^2}{bn} - \frac{y_{...}^2}{abn} \tag{13-6}$$

$$SS_B = \sum_{j=1}^{b} \frac{y_{.j.}^2}{an} - \frac{y_{...}^2}{abn} \tag{13-7}$$

$$SS_{AB} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{y_{ij.}^2}{n} - \frac{y_{...}^2}{abn} - SS_A - SS_B \tag{13-8}$$

$$SS_E = SS_T - SS_{AB} - SS_A - SS_B \tag{13-9}$$

Es común presentar los cálculos en una tabla de análisis de varianza, como la tabla 13-4.

Tabla 13-4 Tabla del análisis de varianza para un modelo factorial de dos factores con efectos fijos

Fuente de variación	Suma de cuadros	Grados de libertad	Cuadro medio	F_0
Tratamientos A	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$\frac{MS_A}{MS_E}$
Tratamientos B	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$\frac{MS_B}{MS_E}$
Interacción	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$		
Total	SS_T	$abn-1$	$MS_E = \frac{SS_E}{ab(n-1)}$	

EJEMPLO 13-4

La pintura tapaporo de aviones se aplica en superficies de aluminio utilizando dos métodos: por inmersión y por aspersión. El objeto de la pintura tapaporo es mejorar la adherencia de la pintura, y en algunas partes puede aplicarse utilizando cualquiera de los dos métodos. Al grupo de ingenieros responsable del proceso de esta operación le interesa saber si tres pinturas tapaporo diferentes difieren en sus propiedades de adherencia. Se realizó un experimento factorial para investigar el efecto que tienen el tipo de pintura tapaporo y el método de aplicación sobre la adherencia de la pintura. Se pintaron tres ejemplares de prueba con cada pintura tapaporo utilizando cada uno de los dos métodos de aplicación, se aplicó la pintura final, y se midió la fuerza de adherencia. En la tabla 13-5 se muestran los datos del experimento. Los números encerrados en círculos son los totales de las celdas $y_{ij\cdot}$. Las sumas de cuadrados que se necesitan para llevar a cabo el análisis de varianza se calculan de la siguiente manera:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{\cdots}^2}{abn}$$

$$= (4.0)^2 + (4.5)^2 + \cdots + (5.0)^2 - \frac{(89.8)^2}{18} = 10.72$$

$$SS_{\text{tipos}} = \sum_{i=1}^{a} \frac{y_{i\cdot\cdot}^2}{bn} - \frac{y_{\cdots}^2}{abn}$$

$$= \frac{(28.7)^2 + (34.1)^2 + (27.0)^2}{6} - \frac{(89.8)^2}{18} = 4.58$$

$$SS_{\text{métodos}} = \sum_{j=1}^{b} \frac{y_{\cdot j\cdot}^2}{an} - \frac{y_{\cdots}^2}{abn}$$

$$= \frac{(40.2)^2 + (49.6)^2}{9} - \frac{(89.8)^2}{18} = 4.91$$

$$SS_{\text{interacción}} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij\cdot}^2}{n} - \frac{y_{\cdots}^2}{abn} - SS_{\text{tipos}} - SS_{\text{métodos}}$$

$$= \frac{(12.8)^2 + (15.9)^2 + (11.5)^2 + (15.9)^2 + (18.2)^2 + (15.5)^2}{3}$$

$$- \frac{(89.8)^2}{18} - 4.58 - 4.91 = 0.24$$

Tabla 13-5 Datos de la fuerza de adherencia del ejemplo 13-4

Tipo de tapaporos	Inmersión		Aspersión		$y_{i\cdot\cdot}$
1	4.0, 4.5, 4.3	12.8	5.4, 4.9, 5.6	15.9	28.7
2	5.6, 4.9, 5.4	15.9	5.8, 6.1, 6.3	18.2	34.1
3	3.8, 3.7, 4.0	11.5	5.5, 5.0, 5.0	15.5	27.0
$y_{\cdot j\cdot}$	40.2		49.6		$89.8 = y_{\cdots}$

Tabla 13-6 Análisis de varianza del ejemplo 13-4

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Tipo de pintura tapaporo	4.58	2	2.29	28.63	2.7 × E-5
Métodos de aplicación	4.91	1	4.91	61.38	5.0 × E-7
Interacción	0.24	2	0.12	1.50	0.2621
Error	0.99	12	0.08		
Total	10.72	17			

y

$$SS_E = SS_T - SS_{\text{tipos}} - SS_{\text{métodos}} - SS_{\text{interacción}}$$
$$= 10.72 - 4.58 - 4.91 - 0.24 = 0.99$$

En la tabla 13-6 se resume el análisis de varianza. El experimentador ha decido usar $\alpha = 0.05$. Puesto que $f_{0.05,\,2,\,12} = 3.89$ y $f_{0.05,\,1,\,12} = 4.75$, se concluye que los efectos principales del tipo de pintura tapaporo y del método de aplicación afectan la fuerza de adherencia. Además, puesto que $1.5 < f_{0.05,\,2,\,12}$, no hay indicios de interacción entre estos factores. En la última columna de la tabla 13-6 se muestra el valor P para cada cociente F. Obsérvese que los valores P de los dos estadísticos de prueba para los efectos principales son considerablemente menores que 0.05, mientras que el valor P para el estadístico de prueba de la interacción es mayor que 0.05.

En la figura 13-10 se presenta una gráfica de los promedios de la fuerza de adherencia de las celdas $\{\bar{y}_{ij\cdot}\}$ contra los niveles del tipo de pintura tapaporo para cada método de aplicación. La conclusión de que no hay interacción es evidente en esta gráfica, ya que las dos líneas son prácticamente paralelas. Además, puesto que una respuesta grande indica una fuerza de adherencia mayor, se concluye que la aspersión es el mejor método de aplicación y que la pintura tapaporo tipo 2 es la más efectiva.

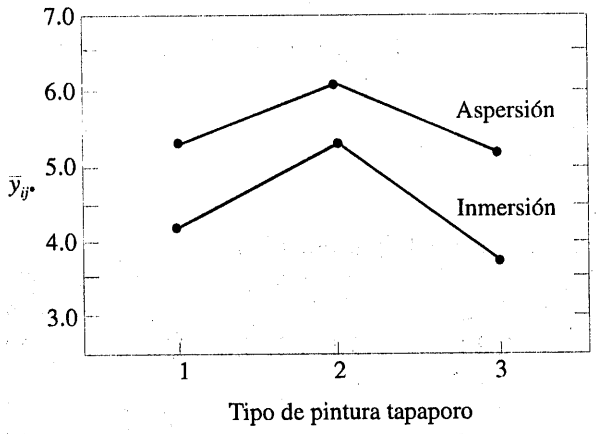


Figura 13-10 Gráfica de la fuerza de adherencia promedio contra los tipos de pintura tapaporo para ambos métodos de aplicación.

Pruebas de medias individuales

Cuando ambos factores son fijos, pueden hacerse comparaciones entre las medias individuales de cualquiera de los dos factores utilizando el método gráfico que se describió en el capítulo 12, o cualquier técnica formal de comparación múltiple como el método de la LSD de Fisher. Cuando no hay interacción, estas comparaciones pueden hacerse utilizando los promedios de los renglones $\bar{y}_{i..}$, o bien los promedios de las columnas $\bar{y}_{.j.}$. Sin embargo, cuando la interacción es significativa, las comparaciones entre las medias de un factor (por ejemplo A) pueden ser oscurecidas por la interacción AB. En este caso, podría aplicarse un procedimiento tal como el método de la LSD de Fisher a las medias del factor A, con el factor B fijo en un nivel particular.

13-4.2 Verificación de la adecuación del modelo

Al igual que en los experimentos con un solo factor que se revisaron en el capítulo 12, los residuales de un experimento factorial desempeñan un papel importante para evaluar la adecuación del modelo. Los residuales de un factorial de dos factores son

$$e_{ijk} = y_{ijk} - \bar{y}_{ij}.$$

Es decir, los residuales son tan sólo la diferencia entre las observaciones y los promedios de las celdas correspondientes.

En la tabla 13-7 se presentan los residuales de los datos de la pintura tapaporo para la pintura de aviones del ejemplo 13-4. La gráfica de probabilidad normal de estos residuales se muestra en la figura 13-11. Esta gráfica tiene colas que no caen exactamente sobre la línea recta que pasa por el centro de la gráfica, lo cual indica algunos problemas potenciales con el supuesto de normalidad, aunque la desviación de la normalidad no parece ser muy marcada. En las figuras 13-12 y 13-13 se grafican los residuales contra los niveles de los tipos de pintura tapaporo y los métodos de aplicación, respectivamente. Hay ciertos indicios de que la pintura tapaporo tipo 3 produce una variabilidad ligeramente más baja en la fuerza de adherencia que las otras dos pinturas tapaporo. La gráfica de los residuales contra los valores ajustados de la figura 13-14 no revela ningún patrón inusual o de diagnóstico.

13-4.3 Salida de computadora

En la tabla 13-8 se muestran algunas salidas del análisis de varianza del programa Minitab para el experimento de la pintura tapaporo de aviones del ejemplo 13-4. En la parte superior de la

Tabla 13-7 Residuales del experimento de la pintura tapaporo de aviones del ejemplo 13-4

	Método de aplicación	
Tipo de pintura tapaporo	Inmersión	Aspersión
1	–0.27, 0.23, 0.03	0.10, –0.40, 0.30
2	0.30, –0.40, 0.10	–0.27, 0.03, 0.23
3	–0.03, –0.13, 0.17	0.33, –0.17, –0.17

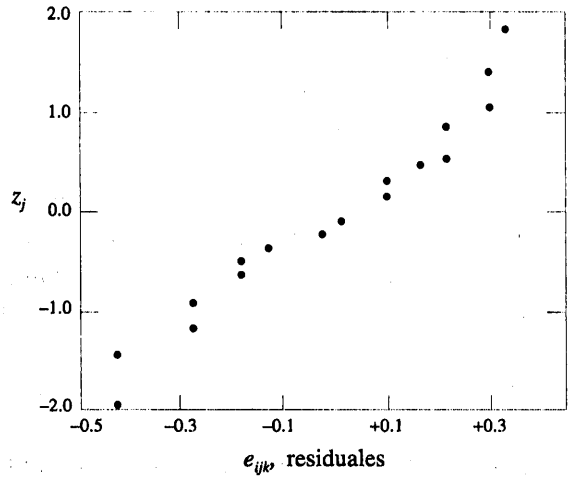


Figura 13-11 Gráfica de probabilidad normal de los residuales del ejemplo 13-4.

tabla se da el nombre del factor e información del nivel, y en la parte inferior de la tabla se presenta el análisis de varianza de la respuesta, la fuerza de adherencia. Los resultados difieren ligeramente de los cálculos manuales que se presentaron en la tabla 13-6 debido a que el software de computadora trabaja con nueve cifras decimales en los cálculos. El programa Minitab también producirá las gráficas de diagnóstico de los residuales presentadas con anterioridad (figuras 13-11 a 13-14) y la gráfica de la interacción de los dos factores de la figura 13-10.

13-4.4 Una observación por celda

En algunos casos en que interviene un experimento factorial con dos factores, puede tenerse una sola réplica; es decir, una sola observación por celda. En esta situación, hay exactamente tantos parámetros en el modelo del análisis de varianza como observaciones, y los grados de libertad

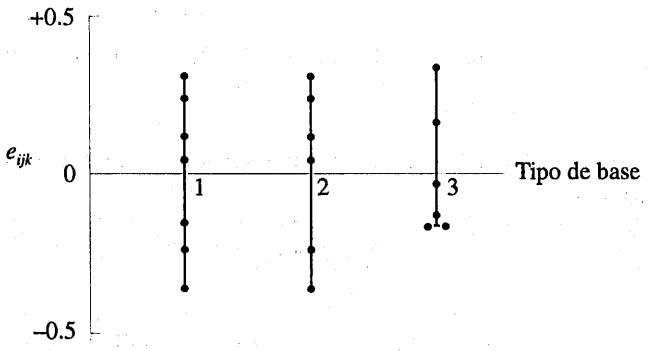


Figura 13-12 Gráfica de los residuales contra el tipo de pintura tapaporo.

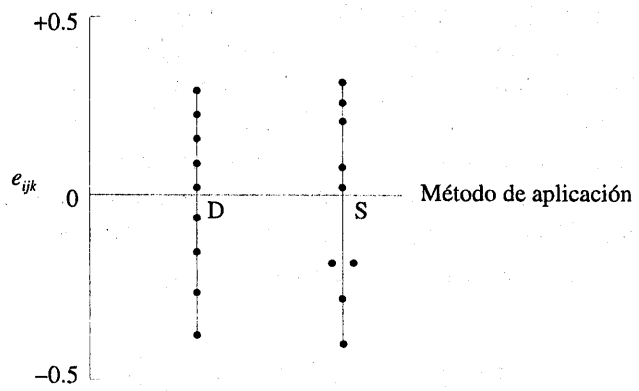


Figura 13-13 Gráfica de los residuales contra el método de aplicación.

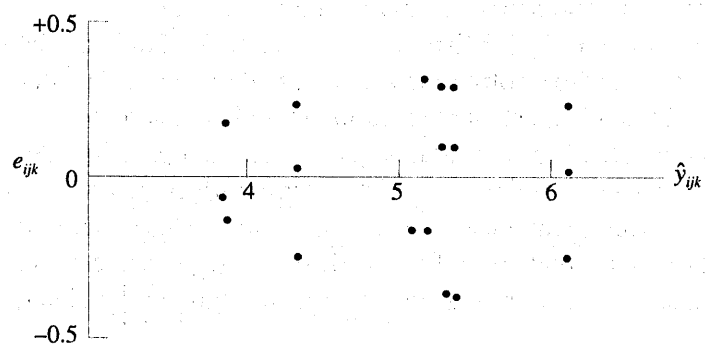


Figura 13-14 Gráfica de los residuales contra los valores predichos $\hat{y}_{ijk}$.

Tabla 13-8 Análisis de varianza del programa Minitab para el ejemplo 13-4

Análisis de varianza (diseños balanceados)

Factor	Tipo	Niveles	Valores		
Pintura tapaporo	fijo	3	1	2	3
Método	fijo	2	Inmersión	Aspersión	

Análisis de varianza de la adherencia

Fuente	GL	SS	MS	F	P
P. tapaporo	2	4.5811	2.2906	27.86	0.000
Método	1	4.9089	4.9089	59.70	0.000
P. tapaporo*Método	2	0.2411	0.1206	1.47	0.269
Error	12	0.9867	0.0822		
Total	17	10.7178			

del error son cero. Por tanto, no es posible probar hipótesis sobre los efectos principales y las interacciones a menos que se hagan supuestos adicionales. Una posible consideración es suponer que el efecto de la interacción se puede omitir y usar el cuadrado medio de la interacción como cuadrado medio de error. Por tanto, el análisis es equivalente al análisis usado en el diseño de bloques aleatorizados. Este supuesto de que no hay interacción puede ser peligroso, y el experimentador deberá examinar cuidadosamente los datos y los residuales en busca de indicios de si está presente o no una interacción. Para mayores detalles, véase Montgomery.

13-4.5 Factores aleatorios

En la sección 12-4 se introdujo el concepto de factor aleatorio. Se trata, desde luego, de una situación en la que el factor de interés tiene un gran número de niveles posibles y el experimentador elige al azar un subconjunto de estos niveles de esa población. Entonces se sacan conclusiones acerca de la población de los niveles del factor.

Pueden presentarse factores aleatorios en experimentos factoriales. Si todos los factores son aleatorios, entonces al modelo del análisis de varianza se le llama **modelo de efectos aleatorios**. Si algunos factores son fijos y otros son aleatorios, al modelo del análisis de varianza se le llama **modelo mixto**. El análisis estadístico de los modelos aleatorios y mixtos es muy similar al de los modelos de efectos fijos estándares que son el tema principal de este capítulo. Las principales diferencias radican en el tipo de hipótesis que se prueban, en la construcción de los estadísticos de la prueba para estas hipótesis, y en la estimación de los parámetros del modelo. En el material complementario del apéndice del final del libro se presenta una breve introducción a estos temas. Para una discusión más amplia, remítase a Montgomery y Neter, Wasserman, Nachtsheim y Kutner.

EJERCICIOS DE LA SECCIÓN 13-4

13-1. En su libro *Diseño y Analisis de Experimentos*, Limusa-Wiley & Sons, D. C. Montgomery presenta los resultados de un experimento en el que interviene una batería de almacenamiento usada en el mecanismo de lanzamiento de un misil tierra-aire para cargar al hombro. Pueden usarse tres tipos de materiales para hacer las placas de la batería. El objetivo es diseñar una batería que se mantenga relativamente sin alteraciones por la temperatura ambiente. La respuesta de salida de la batería es la vida efectiva en horas. Se seleccionan tres niveles de temperatura y se corre un experimento factorial con cuatro réplicas. Los datos son los siguientes:

	Temperatura (°F)					
Material	Baja		Media		Alta	
1	130	155	34	40	20	70
	74	180	80	75	82	58
2	150	188	136	122	25	70
	159	126	106	115	58	45
3	138	110	174	120	96	104
	168	160	150	139	82	60

a) Pruebe las hipótesis apropiadas y saque conclusiones utilizando el análisis de varianza con $\alpha = 0.05$.

b) Analice gráficamente la interacción.

c) Analice los residuales de este experimento.

13-2. Un ingeniero sospecha que en el terminado superficial de ciertas piezas metálicas influyen el tipo de pintura usado y el tiempo de secado. Selecciona tres tiempos de secado —20, 25 y 30 minutos— y usa dos tipos de pintura. Se prueban tres piezas con cada combinación del tipo de pintura y el tiempo de secado. Los datos son los siguientes:

	Tiempo de secado (min)		
Pintura	20	25	30
1	74	73	78
	64	61	85
	50	44	92
2	92	98	66
	86	73	45
	68	88	85

a) Enuncie y pruebe las hipótesis apropiadas utilizando el análisis de varianza con $\alpha = 0.05$.
b) Analice los residuales de este experimento.

13-3. En un artículo de *Industrial Quality Control* (pp. 5-8) se describe un experimento para investigar el efecto de dos factores (tipo de cristal y tipo de fósforo) sobre la brillantez de un cinescopio. La variable de respuesta medida es la corriente (en microamperes) necesaria para obtener un nivel específico de brillantez. Los datos se presentan en la siguiente tabla:

Tipo de	Tipo de fósforo		
cristal	1	2	3
1	280	300	290
	290	310	285
	285	295	290
2	230	260	220
	235	240	225
	240	235	230

a) Enuncie las hipótesis de interés en este experimento.
b) Pruebe las hipótesis anteriores y saque conclusiones utilizando el análisis de varianza con $\alpha = 0.05$.
c) Analice los residuales de este experimento.

13-4. Se condujo un experimento para determinar si la temperatura del fuego o la posición en el horno afectan la densidad de endurecimiento de un ánodo de carbono. Los datos son los siguientes:

	Temperatura (°C)		
Posición	800	825	850
1	570	1 063	565
	565	1 080	510
	583	1 043	590
2	528	988	526
	547	1 026	538
	521	1 004	532

a) Enuncie las hipótesis de interés.
b) Pruebe las hipótesis anteriores utilizando el análisis de varianza con $\alpha = 0.05$. ¿A qué conclusiones se llega?
c) Analice los residuales de este experimento.

13-5. Continuación del ejercicio 13-4. Utilizando el método de la LSD de Fisher, investigar las diferencias entre la media de la densidad del endurecimiento de los ánodos en los tres diferentes niveles de temperatura del ejercicio 13-4. Use $\alpha = 0.05$.

13-6. Johnson y Leone (*Statistics and Experimental Design in Engineering and the Physical Sciences*, John Wiley) describen un experimento que se conduce para investigar la torcedura de placas de cobre. Los dos factores estudiados fueron la temperatura y el contenido de cobre de las placas. La variable de respuesta es la cantidad de deformación. Los datos son los siguientes:

Temperatura (°C)	Contenido de cobre (%) 40	60	80	100
50	17, 20	16, 21	24, 22	28, 27
75	12, 9	18, 13	17, 12	27, 31
100	16, 12	18, 21	25, 23	30, 23
125	21, 17	23, 21	23, 22	29, 31

a) ¿Hay algún indicio de que alguno de los dos factores afecta la cantidad de torcedura? ¿Hay alguna interacción entre los factores? Use $\alpha = 0.05$.

b) Analice los residuales de este experimento.

c) Grafique la torcedura promedio para cada nivel del contenido de cobre y compare los niveles utilizando el método de la LSD de Fisher. Describa las diferencias en los efectos de los diferentes niveles del contenido de cobre sobre la torcedura. Si es deseable una tocedura reducida, ¿qué nivel del contenido de cobre se especificaría?

d) Suponga que no es sencillo controlar la temperatura en el medio ambiente don de van a usarse las placas de cobre. ¿Esto modifica la respuesta dada en el inciso c)?

13-7. Continuación del ejercicio 13-6. Aplique el método gráfico para encontrar las diferencias entre las medias con los datos de la deformación promedio en cada nivel del contenido de cobre del ejercicio 13-6. ¿Se llega a la misma conclusión que en el inciso c) de ese ejercicio, donde se usó el método de la LSD de Fisher?

13-8. En un artículo de *Journal of Testing and Evaluation* (Vol. 16, No. 6, pp. 508-515) se investigaron los efectos de la frecuencia de carga cíclica y de las condiciones ambientales sobre el crecimiento de las fisuras por fatiga con un esfuerzo constante de 22 MPa para un material particular. Los datos del experimento se presentan a continuación. La variable de respuesta es la rapidez de crecimiento de las fisuras por fatiga.

		Medio Ambiente: Aire	H_2O	H_2O Salada
Frecuencia	10	2.29	2.06	1.90
		2.47	2.05	1.93
		2.48	2.23	1.75
		2.12	2.03	2.06
	1	2.65	3.20	3.10
		2.68	3.18	3.24
		2.06	3.96	3.98
		2.38	3.64	3.24
	0.1	2.24	11.00	9.96
		2.71	11.00	10.01
		2.81	9.06	9.36
		2.08	11.30	10.40

a) ¿Hay indicios de que alguno de los dos factores afecta la rapidez de crecimiento de las fisuras por fatiga? ¿Hay indicios de interacción? Use $\alpha = 0.05$.

b) Analice los residuales de este experimento.

c) Repita el análisis del inciso a) usando $ln(y)$ como variable de respuesta. Analice los residuales de esta nueva variable de respuesta y comente los resultados.

13-9. En un artículo de *IEEE Transactions on Electron Devices* se describe un estudio de los efectos de dos variables —el dopado del polisilicio y las condiciones del revenido (tiempo y temperatura)— sobre la corriente fundamental de un transistor bipolar. Los datos de este experimento se muestran al principio de la página siguiente.

a) ¿Hay evidencia que apoye la afirmación de que el dopado del polisilicio o las condiciones del revenido afectan la corriente fundamental? ¿Hay interacción entre estas variables? Use $\alpha = 0.05$.

b) Analice gráficamente la interacción.

c) Analice los residuales de este experimento.

13-10. Considérese el experimento descrito en el ejercicio 13-9.

		Revenido (temperatura/tiempo)				
		900/60	900/180	950/60	1 000/15	1 000/30
Dopado del polisilicio	1×10^{20}	4.40 4.60	8.30 8.90	10.15 10.20	10.29 10.30	11.01 10.58
	2×10^{20}	3.20 3.50	7.81 7.75	9.38 10.02	10.19 10.10	10.81 10.60

a) Use el método gráfico que se describió en el capítulo 12 para aislar los efectos de las condiciones del revenido sobre la corriente fundamental.

b) Repita el inciso a) usando el método de la LSD de Fisher con $\alpha = 0.05$.

c) ¿Ambos métodos llevan a las mismas conclusiones?

13-5 EXPERIMENTOS FACTORIALES GENERALES

En muchos experimentos intervienen más de dos factores. En esta sección se introduce el caso en que hay a niveles del factor A, b niveles del factor B, c niveles del factor C, etc., ordenados en un experimento factorial. En general, habrá en total $abc \cdots n$ observaciones, si hay n repeticiones del experimento completo.

Por ejemplo, considérese el experimento con tres factores que tiene el modelo fundamental

$$Y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk}$$

$$+ (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, c \\ l = 1, 2, \ldots, n \end{cases} \tag{13-10}$$

En la tabla 13-9 se muestra el análisis de varianza, suponiendo que A, B y C son factores fijos. Obsérvese que debe haber al menos dos repeticiones ($n \geq 2$) para calcular la suma de cuadrados del error. La prueba F de los efectos principales y las interacciones se sigue directamente de los cuadrados medios esperados. Estos cocientes siguen una distribución F bajo las respectivas hipótesis nulas.

Las fórmulas para calcular las sumas de cuadrados de la tabla 13-9 se obtienen con facilidad. La suma total de cuadrados es

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} y_{ijkl}^2 - \frac{y_{\ldots}^2}{abcn} \tag{13-11}$$

La suma de cuadrados de los efectos principales se calcula a partir de los totales de los factores $A(y_{i\ldots})$, $B(y_{.j.})$ y $C(y_{..k})$ como sigue:

$$SS_A = \sum_{i=1}^{a} \frac{y_{i\ldots}^2}{bcn} - \frac{y_{\ldots}^2}{abcn} \tag{13-12}$$

Tabla 13-9 Tabla del análisis de varianza del modelo de tres factores con efectos fijos

Fuentes de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	Cuadrados medios esperados	F_0
A	SS_A	$a-1$	MS_A	$\sigma^2 + \frac{bcn\sum\tau_i^2}{a-1}$	$\frac{MS_A}{MS_E}$
B	SS_B	$b-1$	MS_B	$\sigma^2 + \frac{acn\sum\beta_j^2}{b-1}$	$\frac{MS_B}{MS_E}$
C	SS_C	$c-1$	MS_C	$\sigma^2 + \frac{abn\sum\gamma_k^2}{c-1}$	$\frac{MS_C}{MS_E}$
AB	SS_{AB}	$(a-1)(b-1)$	MS_{AB}	$\sigma^2 + \frac{cn\sum\sum(\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
AC	SS_{AC}	$(a-1)(c-1)$	MS_{AC}	$\sigma^2 + \frac{bn\sum\sum(\tau\gamma)_{ik}^2}{(a-1)(c-1)}$	$\frac{MS_{AC}}{MS_E}$
BC	SS_{BC}	$(b-1)(c-1)$	MS_{BC}	$\sigma^2 + \frac{an\sum\sum(\beta\gamma)_{jk}^2}{(b-1)(c-1)}$	$\frac{MS_{BC}}{MS_E}$
ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	MS_{ABC}	$\sigma^2 + \frac{n\sum\sum\sum(\tau\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)}$	$\frac{MS_{ABC}}{MS_E}$
Error	SS_E	$abc(n-1)$	MS_E	σ^2	
Total	SS_T	$abcn-1$			

$$SS_B = \sum_{j=1}^{b} \frac{y_{\cdot j\cdot\cdot}^2}{acn} - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} \tag{13-13}$$

$$SS_C = \sum_{k=1}^{c} \frac{y_{\cdot\cdot k\cdot}^2}{abn} - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} \tag{13-14}$$

Para calcular las sumas de cuadrados de las interacciones de dos factores se necesitan los totales de las celdas $A \times B$, $A \times C$ y $B \times C$. Puede ser conveniente plegar la tabla de los datos originales en tres tablas con dos factores a fin de calcular estos totales. Las sumas de cuadrados son

$$SS_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij\cdot\cdot}^2}{cn} - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} - SS_A - SS_B \tag{13-15}$$

$$SS_{AC} = \sum_{i=1}^{a}\sum_{k=1}^{c} \frac{y_{i\cdot k\cdot}^2}{bn} - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} - SS_A - SS_C \tag{13-16}$$

$$SS_{BC} = \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{\cdot jk\cdot}^2}{an} - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} - SS_B - SS_C \tag{13-17}$$

La suma de cuadrados de la interacción de los tres factores se calcula a partir de los totales de las celdas de los tres factores $\{y_{ijk\cdot}\}$ como

$$SS_{ABC} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \frac{y_{ijk\cdot}^2}{n} - \frac{y_{\cdots\cdot}^2}{abcn} - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC} \tag{13-18}$$

La suma de cuadrados del error puede encontrarse restando la suma de los cuadrados de cada efecto principal y de la interacción de la suma total de cuadrados.

EJEMPLO 13-5

Un ingeniero mecánico estudia la rugosidad superficial de una pieza producida en una operación de corte de metal. Son de interés tres factores: la rapidez de alimentación (A), la profundidad del corte (B) y el ángulo de la herramienta (C). A los tres factores se les ha asignado dos niveles, y se corren dos réplicas de un diseño factorial. En la tabla 13-10 se presentan los datos codificados. Los totales de las celdas con los tres factores $\{y_{ijk\cdot}\}$ están encerrados en un círculo en esta tabla.

Tabla 13-10 Datos codificados de la rugosidad superficial del ejemplo 13-5

Rapidez de alimentación (A)	Profundidad del corte (B): 0.025 pulgadas, Ángulo de la herramienta (C): 15°	0.025 pulgadas, 25°	0.04 pulgadas, 15°	0.04 pulgadas, 25°	$y_{i\cdots}$
20 pulg/min	9 7 16	11 10 21	9 11 20	10 8 18	75
30 pulg/min	10 12 22	10 13 23	12 15 27	16 14 30	102
Totales $B \times C$ $y_{\cdot jk\cdot}$	38	44	47	48	$177 = y_{\cdots\cdot}$

Totales $A \times B$ $y_{ij\cdot\cdot}$

A \ B	0.025	0.040
20	37	38
30	45	57
$y_{\cdot j\cdot\cdot}$	82	95

Totales $A \times C$ $y_{i\cdot k\cdot}$

A \ C	15	25
20	36	39
30	49	53
$y_{\cdot\cdot k\cdot}$	85	92

Las sumas de cuadrados se calculan de la siguiente manera, utilizando las ecuaciones 13-11 a 13-18:

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\sum_{l=1}^{2} y_{ijkl}^2 - \frac{y_{....}^2}{abcn} = 2\,051 - \frac{(177)^2}{16} = 92.9375$$

$$SS_A = \sum_{i=1}^{2} \frac{y_{i\cdots}^2}{bcn} - \frac{y_{....}^2}{abcn}$$

$$= \frac{(75)^2 + (102)^2}{8} - \frac{(177)^2}{16} = 45.5625$$

$$SS_B = \sum_{j=1}^{2} \frac{y_{\cdot j\cdot\cdot}^2}{acn} - \frac{y_{....}^2}{abcn}$$

$$= \frac{(82)^2 + (95)^2}{8} - \frac{(177)^2}{16} = 10.5625$$

$$SS_C = \sum_{k=1}^{2} \frac{y_{\cdot\cdot k\cdot}^2}{abn} - \frac{y_{....}^2}{abcn}$$

$$= \frac{(85)^2 + (92)^2}{8} - \frac{(177)^2}{16} = 3.0625$$

$$SS_{AB} = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{y_{ij\cdot\cdot}^2}{cn} - \frac{y_{....}^2}{abcn} - SS_A - SS_B$$

$$= \frac{(37)^2 + (38)^2 + (45)^2 + (57)^2}{4} - \frac{(177)^2}{16} - 45.5625 - 10.5625$$

$$= 7.5625$$

$$SS_{AC} = \sum_{i=1}^{2}\sum_{k=1}^{2} \frac{y_{i\cdot k\cdot}^2}{bn} - \frac{y_{....}^2}{abcn} - SS_A - SS_C$$

$$= \frac{(36)^2 + (39)^2 + (49)^2 + (53)^2}{4} - \frac{(177)^2}{16} - 45.5625 - 3.0625$$

$$= 0.0625$$

$$SS_{BC} = \sum_{j=1}^{2}\sum_{k=1}^{2} \frac{y_{\cdot jk\cdot}^2}{an} - \frac{y_{....}^2}{abcn} - SS_B - SS_C$$

$$= \frac{(38)^2 + (44)^2 + (47)^2 + (48)^2}{4} - \frac{(177)^2}{16} - 10.5625 - 3.0625$$

$$= 1.5625$$

$$SS_{ABC} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\frac{y_{ijk\cdot}^2}{n} - \frac{y_{\cdots\cdot}^2}{abcn} - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC}$$
$$= \frac{(16)^2+(21)^2+(20)^2+(18)^2+(22)^2+(23)^2+(27)^2+(30)^2}{2}$$
$$-\frac{(177)^2}{16} - 45.5625 - 10.5625 - 3.0625 - 7.5625 - 0.0625 - 1.5625$$
$$= 73.4375 - 45.5625 - 10.5625 - 3.0625 - 7.5625 - 0.0625 - 1.5625$$
$$= 5.0625$$

$$SS_E = SS_T - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC} - SS_{ABC}$$
$$= 92.9375 - 45.5625 - 10.5625 - 3.0625 - 7.5625 - 0.0625$$
$$-1.5625 - 5.0625$$
$$= 19.5000$$

El análisis de varianza se resume en la tabla 13-11. Los cocientes F de los tres efectos principales y las interacciones se forman dividiendo el cuadrado medio del efecto de interés por el cuadrado medio de error. Puesto que el experimentador a seleccionado $\alpha = 0.05$, el valor crítico de cada uno de estos cocientes F es $f_{0.05,\,1,\,8} = 5.32$. De manera alternativa, podría usarse el enfoque del valor P. Los valores P de todos los estadísticos de la prueba se muestran en la última columna de la tabla 13-11. La inspección de estos valores P es reveladora. Hay un mar-

Tabla 13-11 Análisis de varianza del programa Minitab para el ejemplo 13-5

Análisis de varianza (diseños balanceados)

Factor	Tipo	Niveles	Valores	
Alimentación	fijo	2	20	30
Profundidad	fijo	2	0.025	0.040
Ángulo	fijo	2	15	25

Análisis de varianza de la rugosidad

Fuente	GL	SS	MS	F	P
Alimentación	1	45.563	45.563	18.69	0.003
Profundidad	1	10.563	10.563	4.33	0.071
Ángulo	1	3.063	3.063	1.26	0.295
Alimentación*profundidad	1	7.563	7.563	3.10	0.116
Alimentación*ángulo	1	0.062	0.062	0.03	0.877
Profundidad*ángulo	1	1.563	1.563	0.64	0.446
Alimentación*profundidad*ángulo	1	5.062	5.062	2.08	0.188
Error	8	19.500	2.437		
Total	15	92.938			

Tabla 13-12 Análisis de varianza del ejemplo 13-5

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Rapidez de alimentación (A)	45.5625	1	45.5625	18.69	0.0025
Profundidad del corte (B)	10.5625	1	10.5625	4.33	0.0710
Ángulo de la herramienta (C)	3.0625	1	3.0625	1.26	0.2942
AB	7.5625	1	7.5625	3.10	0.1163
AC	0.0625	1	0.0625	0.03	0.8668
BC	1.5625	1	1.5625	0.64	0.4468
ABC	5.0625	1	5.0625	2.08	0.1872
Error	19.5000	8	2.4375		
Total	92.9375	15			

cado efecto principal de la rapidez de alimentación, ya que el cociente F está muy adentro de la región crítica. Sin embargo, hay indicios parciales de un efecto debido a la profundidad del corte, ya que $P = 0.0710$ no es mucho mayor que $\alpha = 0.05$. El siguiente efecto más grande es la interacción AB o rapidez de alimentación × profundidad del corte. Muy probablemente, tanto la rapidez de alimentación como la profundidad del corte son variables importantes del proceso.

El análisis de varianza de un experimento factorial general requiere laboriosos cálculos aritméticos. En la práctica, se usan casi exclusivamente paquetes de software de computadora para el análisis. En la tabla 13-12 se presenta la salida del programa Minitab para el ejemplo 13-5. Los resultados de computadora concuerdan bastante bien con los cálculos manuales que se presentan en la tabla 13-11.

Evidentemente, los experimentos factoriales con tres o más factores pueden requerir muchas corridas experimentales, en particular si alguno de los factores tiene varios niveles (más de dos). Este punto de vista lleva a una clase de diseño factorial en el que todos los factores tienen dos niveles. Es fácil establecer y analizar estos diseños y, como se verá, pueden usarse como base para muchos otros diseños experimentales útiles.

EJERCICIOS DE LA SECCIÓN 13-5

13-11. Se investigan el porcentaje de la concentración de madera dura en la pulpa cruda, la libertad de orientación de la fibra o lof, y el tiempo de cocción de la pulpa en cuanto a sus efectos sobre la resistencia del papel. En la siguiente tabla se muestran los datos de un experimento factorial con tres factores.

Porcentaje de la concentración de madera dura	1.5 horas de tiempo de cocción, lof 350	1.5 horas, lof 500	1.5 horas, lof 650	2.0 horas de tiempo de cocción, lof 350	2.0 horas, lof 500	2.0 horas, lof 650
10	96.6	97.7	99.4	98.4	99.6	100.6
	96.0	96.0	99.8	98.6	100.4	100.9
15	98.5	96.0	98.4	97.5	98.7	99.6
	97.2	96.9	97.6	98.1	96.0	99.0
20	97.5	95.6	97.4	97.6	97.0	98.5
	96.6	96.2	98.1	98.4	97.8	99.8

a) Analice los datos usando el análisis de varianza bajo el supuesto de que todos los factores son fijos. Use $\alpha = 0.05$.

b) Encuentre los valores de P de los cocientes F del inciso a).

c) Los residuales se encuentran con $e_{ijkl} = y_{ijkl} - \bar{y}_{ijk\cdot\cdot}$. Analice gráficamente los residuales de este experimento.

13-12. El departamento de control de calidad de una planta de acabados textiles estudia los efectos de varios factores sobre el teñido de una tela combinada de algodón y fibra sintética que se usa para hacer camisas. Se seleccionan tres operadores, tres duraciones del ciclo y dos temperaturas, y tres ejemplares de prueba pequeños de tela se tiñeron bajo cada conjunto de condiciones. La tela terminada se comparó con un patrón y se asignó una puntuación numérica. Los resultados se presentan en la tabla siguiente.

a) Enuncie y pruebe las hipótesis apropiadas usando el análisis de varianza con $\alpha = 0.05$.

b) Los residuales pueden obtenerse con $e_{ijkl} = y_{ijkl} - y_{ijk\cdot\cdot}$; analice gráficamente los residuales de este experimento.

	Temperatura					
	300°			350°		
	Operador			Operador		
Duración del ciclo	1	2	3	1	2	3
	23	27	31	24	38	34
40	24	28	32	23	36	36
	25	26	28	28	35	39
	36	34	33	37	34	34
50	35	38	34	39	38	36
	36	39	35	35	36	31
	28	35	26	26	36	28
60	24	35	27	29	37	26
	27	34	25	25	34	34

13-6 DISEÑO FACTORIAL 2^k

Es frecuente el uso de diseños factoriales en experimentos en los que intervienen varios factores cuando es necesario estudiar el efecto conjunto de los factores sobre una respuesta. Sin embargo, varios casos especiales del diseño factorial general son importantes debido a su uso generalizado en el trabajo de investigación y debido a que constituyen la base de otros diseños que tienen un importante valor práctico.

El caso especial más importante es el de k factores con sólo dos niveles cada uno. Estos niveles pueden ser cuantitativos, como dos valores de temperatura, presión o tiempo; o pueden ser cualitativos, como dos máquinas, dos operadores, lo niveles "alto" y "bajo" de un factor o quizá la presencia o ausencia de un factor. Una réplica completa de un diseño como éste requiere $2 \times 2 \times \cdots \times 2 = 2^k$ observaciones y se le llama **diseño factorial 2^k**.

El diseño 2^k es de particular utilidad en las etapas iniciales del trabajo experimental, cuando probablemente se estén investigando muchos factores. Proporciona el menor número de corri-

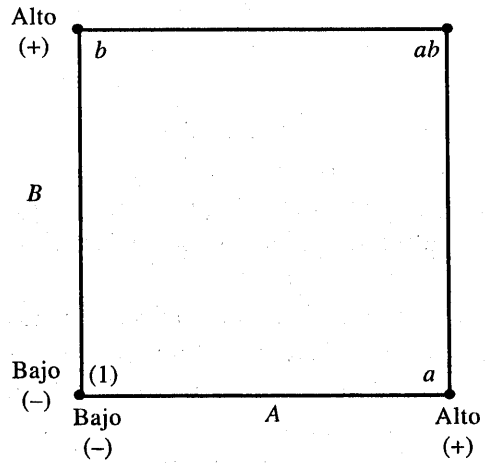
das con las que pueden estudiarse k factores en un diseño factorial completo. Debido a que sólo hay dos niveles para cada factor, debe suponerse que la respuesta es aproximadamente lineal en el rango escogido para los niveles de los factores.

13-6.1 Diseño 2^2

El diseño 2^k más simple es el 2^2; es decir, dos factores A y B, cada uno con dos niveles. Por lo general estos niveles se consideran los niveles bajo y alto del factor. En la figura 13-15 se muestra el diseño 2^2. Obsérvese que el diseño puede representarse geométricamente como un cuadrado en el que las $2^2 = 4$ corridas, o combinaciones de los tratamientos, forman los vértices del cuadrado. En el diseño 2^2 es común denotar los niveles bajo y alto de los factores A y B por los signos – y +, respectivamente. A esto se le llama en ocasiones la **notación geométrica** del diseño.

Se usa una notación especial para etiquetar a las combinaciones de los tratamientos. En general, una combinación de tratamientos se representa por una serie de letras minúsculas. Si una letra está presente, entonces el factor correspondiente se corre con el nivel alto en esa combinación de tratamientos; si está ausente, el factor se corre en el nivel bajo. Por ejemplo, la combinación de tratamientos a indica que el factor A está en el nivel alto y que el factor B está en el nivel bajo. La combinación de tratamientos con ambos factores en el nivel bajo se representa por (1). Esta notación se usa en todas las series de diseños 2^k. Por ejemplo, la combinación de tratamientos en un diseño 2^4 con A y C en el nivel alto y B y D en el nivel bajo se denota por ac.

Los efectos de interés en el diseño 2^2 son los efectos principales de A y B y la interacción de dos factores AB. Sea que las letras (1), a, b y ab representen también los totales de todas las n observaciones tomadas en estos puntos del diseño. Es sencillo estimar los efectos de estos factores. Para estimar el efecto principal de A, se sacaría el promedio de las observaciones de la parte derecha del cuadrado de la figura 13-15 donde A está en el nivel alto, y se le restaría el promedio de las observaciones de la parte izquierda del cuadrado, donde A está en el nivel bajo, o



Tratamiento	A	B
(1)	–	–
a	+	–
b	–	+
ab	+	+

Figura 13-15 El diseño factorial 2^2.

$$
\begin{aligned}
A &= \bar{y}_{A+} - \bar{y}_{A-} \\
&= \frac{a+ab}{2n} - \frac{b+(1)}{2n} \\
&= \frac{1}{2n}\left[a+ab-b-(1)\right]
\end{aligned}
\tag{13-19}
$$

De manera similar, el efecto principal de B se encuentra sacando el promedio de las observaciones de la parte superior del cuadrado, donde B está en el nivel alto, y restando el promedio de las observaciones de la parte inferior del cuadrado, donde B está en el nivel bajo:

$$
\begin{aligned}
B &= \bar{y}_{B+} - \bar{y}_{B-} \\
&= \frac{b+ab}{2n} - \frac{a+(1)}{2n} \\
&= \frac{1}{2n}\left[b+ab-a-(1)\right]
\end{aligned}
\tag{13-20}
$$

Por último, la interacción AB se estima sacando la diferencia de los promedios de las diagonales de la figura 13-15, o

$$
\begin{aligned}
AB &= \frac{ab+(1)}{2n} - \frac{a+b}{2n} \\
&= \frac{1}{2n}\left[ab+(1)-a-b\right]
\end{aligned}
\tag{13-21}
$$

A las cantidades entre corchetes de las ecuaciones 13-19, 13-20 y 13-21 se les llama **contrastes**. Por ejemplo, el contraste A es

$$\text{Contraste}_A = a+ab-b-(1)$$

En estas ecuaciones, los coeficientes de contraste son siempre +1 o −1. Puede usarse una tabla de signos positivos y negativos, como la tabla 13-13, para determinar el signo de cada combina-

Tabla 13-13 Signos de los efectos del diseño 2^2

Combinación de tratamientos	Efecto factorial			
	I	A	B	AB
(1)	+	−	−	+
a	+	+	−	−
b	+	−	+	−
ab	+	+	+	+

ción de tratamientos para un contraste particular. Los encabezados de las columnas de la tabla 13-13 son los efectos principales de A y B, la interacción AB e I, que representa el total. Los títulos de los renglones son las combinaciones de tratamientos. Obsérvese que los signos de la columna AB son el producto de los signos de las columnas A y B. Para generar un contraste a partir de esta tabla, se multiplican los signos de la columna apropiada de la tabla 13-13 por las combinaciones de tratamientos enlistadas en los renglones y se suman. Por ejemplo, $\text{contraste}_{AB} = [(1)] + [-a] + [-b] + [ab] = ab + (1) - a - b$.

Los contrastes se usan para calcular tanto las estimaciones de los efectos como las sumas de cuadrados de A, B y la interacción AB. Las fórmulas de las sumas de los cuadrados son

$$SS_A = \frac{[a + ab - b - (1)]^2}{4n}$$

$$SS_B = \frac{[b + ab - a - (1)]^2}{4n}$$

$$SS_{AB} = \frac{[ab + (1) - a - b]^2}{4n}$$

El análisis de varianza se completa calculando la suma total de cuadrados SS_T (con $4n - 1$ grados de libertad) en la forma acostumbrada y obteniendo la suma de cuadrados del error SS_E [con $4(n - 1)$ grados de libertad] por sustracción.

EJEMPLO 13-6

En un artículo de *AT&T Technical Journal* (Vol. 65, pp. 39-50) se describe la aplicación de diseños factoriales con dos niveles en la fabricación de circuitos integrados. Un paso básico en esta industria es hacer crecer una capa epitaxial sobre obleas de silicio pulidas. Las obleas se montan en un susceptor y se colocan en el interior de un recipiente con forma de campana. Se introducen vapores químicos por medio de boquillas localizadas cerca de la parte superior de la campana. El susceptor se hace girar, y se aplica calor. Estas condiciones se mantienen hasta que la capa epitaxial tiene el espesor suficiente.

En la tabla 13-14 se presentan los resultados de un diseño factorial 2^2 con $n = 4$ réplicas usando los factores A = tiempo de deposición y B = razón del flujo del arsénico. Los dos niveles del tiempo de deposición son – = corto y + = largo, y los dos niveles de la razón de flujo del arsénico son – = 55% y + = 59%. La variable de respuesta es el espesor de la capa epitaxial (μm). Las estimaciones de los efectos pueden encontrarse utilizando las ecuaciones 13-19, 13-20 y 13-21 de la siguiente manera:

$$A = \frac{1}{2n}[a + ab - b - (1)]$$
$$= \frac{1}{2(4)}[59.299 + 59.156 - 55.686 + 56.081] = 0.836$$
$$B = \frac{1}{2n}[b + ab - a - (1)]$$
$$= \frac{1}{2(4)}[55.686 + 59.156 - 56.299 - 56.081] = 0.067$$
$$AB = \frac{1}{2n}[ab + (1) - a - b]$$
$$AB = \frac{1}{2(4)}[59.156 + 56.081 - 59.299 - 55.686] = 0.032$$

Las estimaciones numéricas de los efectos indican que el efecto del tiempo de deposición es grande y que tiene una dirección positiva (el incremento del tiempo de deposición aumenta el espesor), ya que al cambiar el tiempo de deposición de bajo a alto se modifica el espesor medio de la capa epitaxial en 0.836 μm. Los efectos de la razón de flujo del arsénico (B) y de la interacción AB parecen ser reducidos.

La importancia de estos efectos puede confirmarse con el análisis de varianza. Las sumas de cuadrados de A, B y AB se calculan de la siguiente manera:

$$SS_A = \frac{[a + ab - b - (1)]^2}{16}$$
$$= \frac{[6.688]^2}{16}$$
$$= 2.7956$$

Tabla 13-14 El diseño 2^2 para el experimento del proceso epitaxial

Combinación de tratamientos	Factores del diseño			Espesor (μm)				Espesor (μm)	
	A	B	AB					Total	Promedio
(1)	–	–	+	14.037	14.165	13.972	13.907	56.081	14.020
a	+	–	–	14.821	14.757	14.843	14.878	59.299	14.825
b	–	+	–	13.880	13.860	14.032	13.914	55.686	13.922
ab	+	+	+	14.888	14.921	14.415	14.932	59.156	14.789

$$SS_A = \frac{[b+ab-a-(1)]^2}{16} = \frac{[-0.538]^2}{16} = 0.0181$$

$$SS_{AB} = \frac{[ab+(1)-a-b]^2}{16} = \frac{[0.252]^2}{16} = -0.0040$$

$$SS_T = 14.037^2 + \cdots + 14.932^2 - \frac{(56.081+\cdots+59.156)^2}{16} = 3.0672$$

El análisis de varianza se resume en la tabla 13-15 y confirma las conclusiones obtenidas al examinar la magnitud y la dirección de los efectos. El tiempo de deposición es el único factor que afecta significativamente el espesor de la capa epitaxial, y por la dirección de las estimaciones del efecto se sabe que los tiempos de deposición más largos producen capas epitaxiales más gruesas.

Análisis de residuales

Es sencillo obtener los residuales de un diseño 2^k ajustando un modelo de regresión a los datos. Para el experimento del proceso epitaxial, el modelo de regresión es

$$Y = \beta_0 + \beta_1 x_1 + \varepsilon$$

ya que la única variable activa es el tiempo de deposición, el cual se representa por la variable codificada x_1. A los niveles bajo y alto del tiempo de deposición se les asignan los valores $x_1 = -1$ y $x_1 = +1$, respectivamente. El modelo de mínimos cuadrados ajustado es

$$\hat{y} = 14.389 + \left(\frac{0.836}{2}\right) x_1$$

Tabla 13-15 Análisis de varianza para el experimento del proceso epitaxial

Fuente de variación	Suma de cuadrados	Grado de libertad	Cuadrado medio	f_0	Valor P
A (tiempo de deposición)	2.7956	1	2.7956	134.40	7.07 E-8
B (flujo de arsénico)	0.0181	1	0.0181	0.87	0.38
AB	0.0040	1	0.0040	0.19	0.67
Error	0.2495	12	0.0208		
Total	3.0672	15			

donde la ordenada al origen $\hat{\beta}_0$ es el gran promedio de las 16 observaciones ($\bar{y}$) y la pendiente $\hat{\beta}_1$ es la mitad de la estimación del efecto del tiempo de deposición. (El coeficiente de regresión es la mitad de la estimación del efecto debido a que los coeficientes de regresión miden el efecto de un cambio unitario de x_1 sobre la media Y, y la estimación del efecto se basa en un cambio de dos unidades de –1 a +1.)

Este modelo puede usarse para obtener los valores predichos en los cuatro puntos que forman los vértices del cuadrado del diseño. Por ejemplo, considérese el punto con tiempo de deposición bajo ($x_1 = -1$) y razón de flujo del arsénico bajo. El valor predicho es

$$\hat{y}=14.389+\left(\frac{0.836}{2}\right)(-1)=13.971\mu\text{m}$$

y los residuales serían

$$\begin{aligned} e_1 &= 14.037-13.971=0.066 \\ e_2 &= 14.165-13.971=0.194 \\ e_3 &= 13.972-13.971=0.001 \\ e_4 &= 13.907-13.971=-0.064 \end{aligned}$$

Es sencillo verificar que los valores predichos y los residuales restantes son, para el tiempo de deposición bajo ($x_1 = -1$) y la razón de flujo del arsénico alto, $\hat{y}$ = 14.389 + (0.836/2)(–1) = 13.971 μm

$$\begin{aligned} e_5 &= 13.880-13.971=-0.091 \\ e_6 &= 13.860-13.971=-0.111 \\ e_7 &= 14.032-13.971=0.061 \\ e_8 &= 13.914-13.971=-0.057 \end{aligned}$$

para el tiempo de deposición alto ($x_1 = +1$) y la razón de flujo del arsénico bajo, $\hat{y}$ = 14.389 + (0.836/2)(+1) = 14.807 μm

$$\begin{aligned} e_9 &= 14.821-14.807=0.014 \\ e_{10} &= 14.757-14.807=-0.050 \\ e_{11} &= 14.843-14.807=0.036 \\ e_{12} &= 14.878-14.807=0.071 \end{aligned}$$

y para el tiempo de deposición alto ($x_1 = +1$) y la razón de flujo del arsénico alto, $\hat{y}$ = 14.389 + (0.836/2)(+1) = 14.807 μm

$$\begin{aligned} e_{13} &= 14.888-14.807=0.081 \\ e_{14} &= 14.921-14.807=0.114 \\ e_{15} &= 14.415-14.807=-0.392 \\ e_{16} &= 14.932-14.807=0.125 \end{aligned}$$

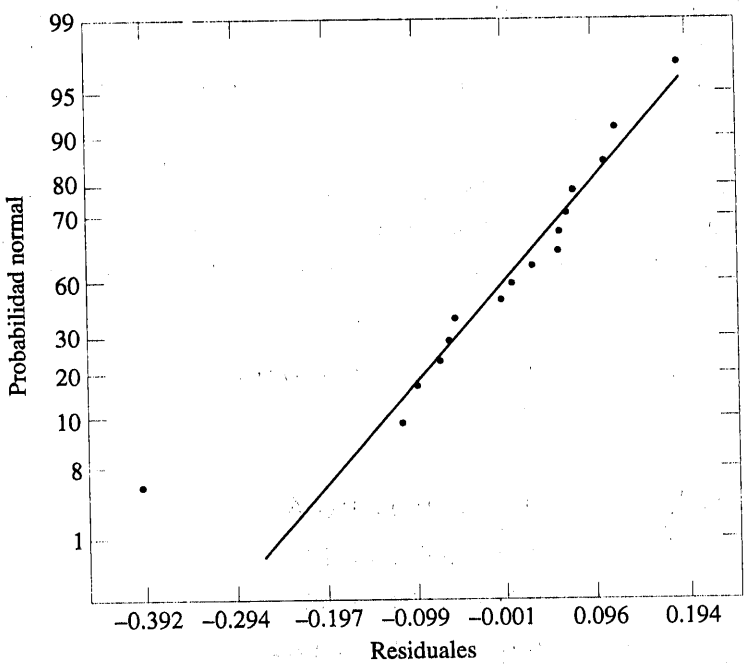


Figura 13-16 Gráfica de probabilidad normal de los residuales del experimento del proceso epitaxial.

En la figura 13-16 se muestra la gráfica de probabilidad normal de estos residuales. La gráfica indica que uno de los residuales, $e_{15} = -0.392$, es un punto atípico. El examen de las cuatro corridas con tiempo de deposición alto y razón de flujo del arsénico alto revela que la observación $y_{15} = 14.415$ es considerablemente menor que las otras tres observaciones para esa combinación de tratamientos. Esto aporta evidencia adicional para la conclusión tentativa de que la observación 15 es un punto atípico. Otra posibilidad es que algunas de las variables del proceso afecten la *variabilidad* del espesor de la capa epitaxial. Si pudiera descubrirse qué variables producen este efecto, quizá sería posible ajustar estas variables en niveles que minimizaran la variabilidad del espesor de la capa epitaxial. Esto podría tener implicaciones importantes en las etapas de manufactura subsecuentes. Las figuras 13-17 y 13-18 son gráficas de los residuales contra el tiempo de deposición y la razón de flujo del arsénico, respectivamente. Además del residual excepcionalmente elevado asociado con y_{15}, no hay evidencia sólida de que ya sea el tiempo de deposición o la razón de flujo del arsénico influya en la variabilidad del espesor de la capa epitaxial.

En la figura 13-19 se muestra la desviación estándar del espesor de la capa epitaxial en las cuatro corridas del diseño 2^2. Estas desviaciones estándar se calcularon usando los datos de la tabla 13-14. Nótese que la desviación estándar de las cuatro observaciones con A y B en el nivel alto son considerablemente mayores que las desviaciones estándar en cualquiera de los otros tres puntos del diseño. La mayor parte de esta diferencia es atribuible a la medición del espesor excepcionalmente bajo asociada con y_{15}. La desviación estándar de las cuatro observaciones con A y B en el nivel bajo también es un tanto mayor que las desviaciones estándar de las dos corridas restantes. Esto podría indicar que otras variables del proceso no incluidas en este experimento pueden afectar la variabilidad del espesor de la capa epitaxial. Podría diseñarse y reali-

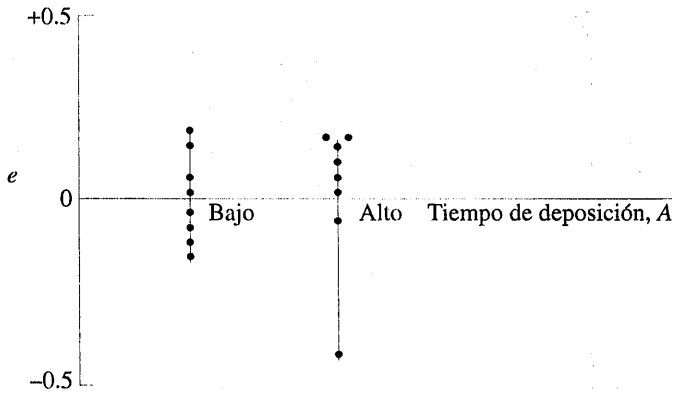


Figura 13-17 Gráfica de los residuales contra el tiempo de deposición.

zarse otro experimento para estudiar esta posibilidad considerando otras variables del proceso. (El artículo original de *AT&T Technical Journal* señala que dos factores adicionales, no considerados en este ejemplo, afectan la variabilidad del proceso.)

13-6.2 Diseño 2^k para $k \geq 3$ factores

El método presentado en la sección anterior para los diseños factoriales con $k = 2$ factores con dos niveles cada uno puede extenderse fácilmente a más de dos factores. Por ejemplo, considérense $k = 3$ factores con dos niveles cada uno. Este diseño es un diseño factorial 2^3, y tiene ocho corridas o combinaciones de tratamientos. Geométricamente, el diseño es un cubo como el que se muestra en la figura 13-20, con las ocho corridas formando los vértices del cubo. Este diseño permite estimar tres efectos principales (A, B y C) junto con tres interacciones de dos factores (AB, AC y BC) y una interacción de tres factores (ABC).

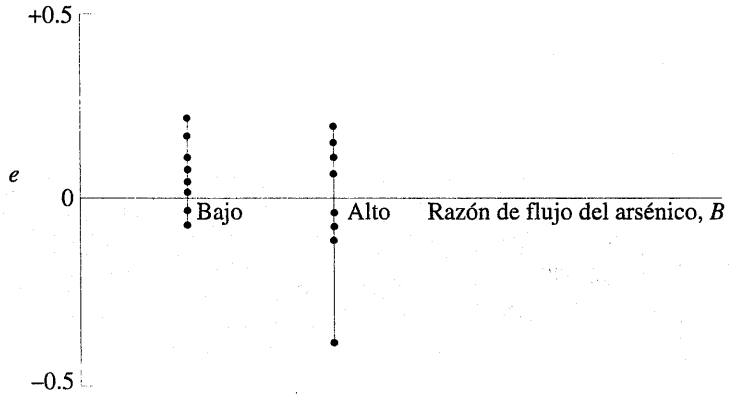


Figura 13-18 Gráfica de los residuales contra la razón de flujo del arsénico.

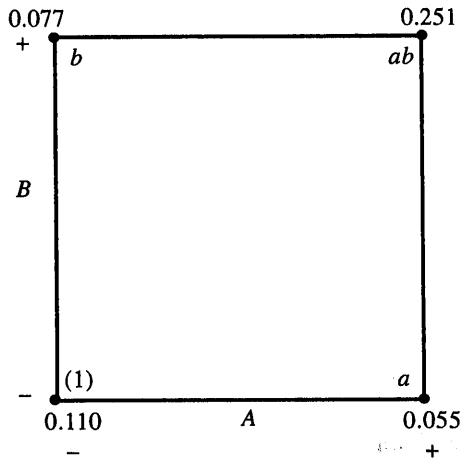


Figura 13-19 La desviación estándar del espesor de la capa epitaxial en las cuatro corridas del diseño 2^2.

Los efectos principales pueden estimarse con facilidad. Recuérdese que las letras minúsculas (1), *a, b, ab, c, ac, bc* y *abc* representan el total de las n réplicas en cada una de las ocho corridas del diseño. Como se observa en la figura 13-21*a*, nótese que el efecto principal de A puede estimarse sacando el promedio de las cuatro combinaciones de tratamientos de la parte derecha del cubo, donde A está en el nivel alto, y restando de esta cantidad el promedio de las cuatro combinaciones de tratamientos de la parte izquierda del cubo, donde A está en el nivel bajo. Se obtiene así

$$\begin{aligned} A &= \bar{y}_{A+} - \bar{y}_{A-} \\ &= \frac{a+ab+ac+abc}{4n} - \frac{(1)+b+c+bc}{4n} \end{aligned}$$

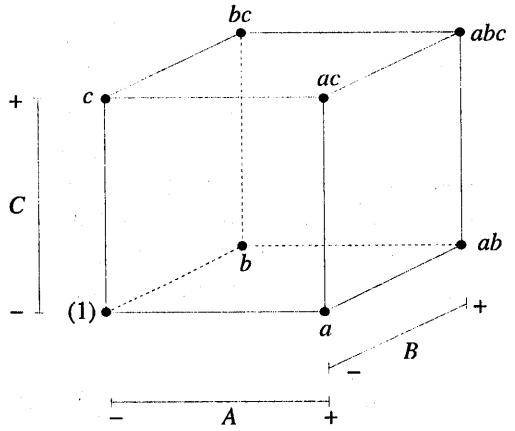


Figura 13-20 El diseño 2^3.

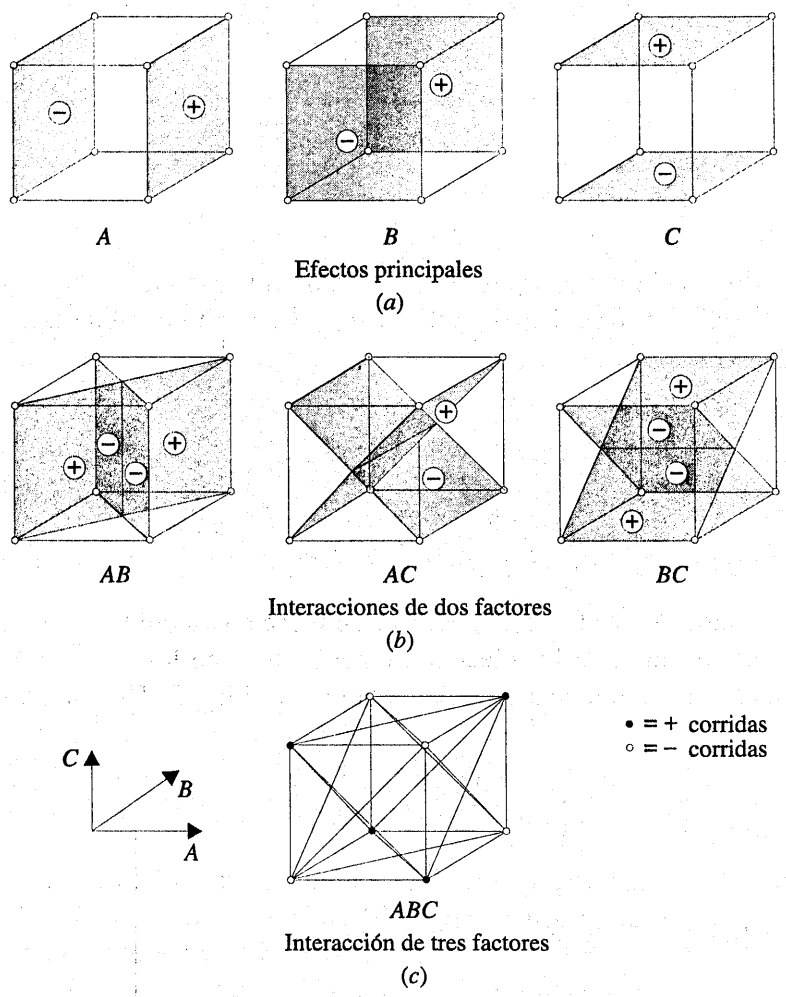


Figura 13-21 Representación geométrica de los contrastes correspondientes a los efectos principales y las interacciones del diseño 2^3. ***a***) Efectos principales. ***b***) Interacciones de dos factores. ***c***) Interacción de tres factores.

Esta ecuación puede arreglarse como

$$A = \frac{1}{4n}\left[a + ab + ac + abc - (1) - b - c - bc\right] \tag{13-22}$$

De manera similar, el efecto de B es la diferencia en los promedios entre las cuatro combinaciones de tratamientos de la cara posterior del cubo (figura 13-21a), y las cuatro combinaciones de la cara frontal. Se obtiene así

$$B = \bar{y}_{B+} - \bar{y}_{B-}$$
$$= \frac{1}{4n}[b + ab + bc + abc - (1) - a - c - ac] \qquad (13\text{-}23)$$

El efecto de C es la diferencia en la respuesta promedio entre las cuatro combinaciones de tratamientos de la cara superior del cubo (figura 13-21a) y las cuatro combinaciones de la cara inferior, es decir,

$$C = \bar{y}_{C+} - \bar{y}_{C-}$$
$$= \frac{1}{4n}[c + ac + bc + abc - (1) - a - b - ab] \qquad (13\text{-}24)$$

Los efectos de la interacción de dos factores pueden calcularse con facilidad. Una medida de la interacción AB es la diferencia entre los efectos promedio de A en los dos niveles de B. Por convención, a la mitad de esta diferencia se le llama la interacción AB. Utilizando símbolos,

B	Efecto promedio de A
Alto (+)	$\dfrac{[(abc - bc) + (ab - b)]}{2n}$
Bajo (–)	$\dfrac{\{(ac - c) + [a - (1)]\}}{2n}$
Diferencia	$\dfrac{[abc - bc + ab - b - ac + c - a + (1)]}{2n}$

Puesto que la interacción AB es la mitad de esta diferencia,

$$AB = \frac{1}{4n}[abc - bc + ab - b - ac + c - a + (1)] \qquad (13\text{-}25)$$

La ecuación 13-25 puede escribirse de la siguiente manera:

$$AB = \frac{abc + ab + c + (1)}{4n} - \frac{bc + b + ac + a}{4n}$$

En esta forma, es fácil ver que la interacción AB es la diferencia de los promedios entre las corridas en dos planos diagonales del cubo de la figura 13-21*b*. Aplicando un razonamiento similar y refiriéndose a la figura 13-21*b*, se encuentra que las interacciones AC y BC son

$$AC = \frac{2}{4n}\left[(1) - a + b - ab - c + ac - bc + abc\right] \tag{13-26}$$

$$AC = \frac{1}{4n}\left[(1) + a - b - ab - c - ac + bc + abc\right] \tag{13-27}$$

La interacción ABC se define como la diferencia promedio entre la interacción AB para los dos diferentes niveles de C. Por tanto,

$$ABC = \frac{2}{4n}\{[abc - bc] - [ac - c] - [ab - b] + [a - (1)]\}$$

o bien

$$ABC = \frac{1}{4n}\left[abc - bc - ac + c - ab + b + a - (1)\right] \tag{13-28}$$

Como anteriormente, la interacción ABC puede considerarse como la diferencia de dos promedios. Si se aíslan las corridas de los dos promedios, éstos definen los vértices de los dos tetraedros que componen el cubo de la figura 13-21*c*.

En las ecuaciones 13-22 a 13-28, las cantidades entre corchetes son los contrastes en las combinaciones de tratamientos. Es posible desarrollar una tabla de signos positivos y negativos a partir de los contrastes, la cual se muestra en la tabla 13-16. Los signos de los efectos principa-

Tabla 13-16 Signos algebraicos para calcular los efectos del diseño 2^3

Combinación de tratamientos	Efecto factorial							
	I	A	B	AB	C	AC	BC	ABC
(1)	+	−	−	+	−	+	+	−
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
ab	+	+	+	+	−	−	−	−
c	+	−	−	+	+	−	−	+
ac	+	+	−	−	+	+	−	−
bc	+	−	+	−	+	−	+	−
abc	+	+	+	+	+	+	+	+

les se determinan asociando un signo positivo con el nivel alto y un signo negativo con el nivel bajo. Una vez que se han establecido los signos de los efectos principales, los signos de las columnas restantes pueden obtenerse multiplicando las columnas precedentes apropiadas, renglón por renglón. Por ejemplo, los signos de la columna AB son los productos de los signos de las columnas A y B en cada renglón. El contraste de cualquier efecto puede obtenerse fácilmente con esta tabla.

La tabla 13-16 posee varias propiedades interesantes:

1. Con excepción de la columna identidad I, cada columna tiene el mismo número de signos positivos y negativos.
2. La suma de los productos de los signos de dos columnas cualesquiera es cero; es decir, las columnas de la tabla son **ortogonales**.
3. Al multiplicar cualquier columna por la columna I, la columna no cambia; es decir, I es un **elemento identidad**.
4. El producto de dos columnas cualesquiera tiene como resultado una columna de la tabla, por ejemplo, $A \times B = AB$, y $AB \times ABC = A^2B^2C = C$, ya que cualquier columna multiplicada por sí misma es la columna identidad.

La estimación de cualquier efecto principal o interacción en un diseño 2^k se determina multiplicando las combinaciones de tratamientos de la primera columna de la tabla por los signos de la columna del efecto principal o columna de interacción correspondiente, sumando el resultado para producir un contraste, y dividiendo después el contraste por la mitad del número total de corridas del experimento. Para cualquier diseño 2^k con n repeticiones, las estimaciones de los efectos se calculan con

$$\text{Efecto} = \frac{\text{Contraste}}{n2^{k-1}} \tag{13-29}$$

y la suma de cuadrados de cualquier efecto es

$$SS = \frac{(\text{Contraste})^2}{n2^k} \tag{13-30}$$

EJEMPLO 13-7

Considérese el experimento de la rugosidad superficial descrito originalmente en el ejemplo 13-5. Se trata de un diseño factorial 2^3 con los factores rapidez de alimentación (A), profundidad del corte (B) y ángulo de la herramienta (C), con $n = 2$ réplicas. En la tabla 13-17 se presentan los datos observados de la rugosidad superficial.

Tabla 13-17 Datos de la rugosidad superficial del ejemplo 13-7

Combinaciones de tratamientos	Factores de diseño A	B	C	Rugosidad superficial	Totales
(1)	−1	−1	−1	9, 7	16
a	1	−1	−1	10, 12	22
b	−1	1	−1	9, 11	20
ab	1	1	−1	12, 15	27
c	−1	−1	1	11, 10	21
ac	1	−1	1	10, 13	23
bc	−1	1	1	10, 8	18
abc	1	1	1	16, 14	30

Los efectos principales pueden estimarse usando las ecuaciones 13-22 a 13-28. El efecto de A, por ejemplo es

$$\begin{aligned} A &= \frac{1}{4n}\left[a+ab+ac+abc-(1)-b-c-bc\right] \\ &= \frac{1}{4(2)}\left[22+27+23+40-16-20-21-18\right] \\ &= \frac{1}{8}\left[27\right] = 3.375 \end{aligned}$$

y la suma de cuadrados de A se encuentra usando la ecuación 13-30:

$$\begin{aligned} SS_A &= \frac{(\text{Contraste}_A)^2}{n2^k} \\ &= \frac{(27)^2}{2(8)} = 45.5625 \end{aligned}$$

Es sencillo verificar que los otros efectos son

$$\begin{aligned} B &= 1.625 \\ C &= 0.875 \\ AB &= 1.375 \\ AC &= 0.125 \\ BC &= -0.625 \\ ABC &= 1.125 \end{aligned}$$

El examen de la magnitud de los efectos revela con claridad que la rapidez de alimentación (factor A) es dominante, seguida por la profundidad del corte (B) y la interacción AB, aun cuan-

Tabla 13-18 Análisis de varianza del experimento del terminado superficial

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
A	45.5625	1	45.5625	18.69	0.0025
B	10.5625	1	10.5625	4.33	0.0709
C	3.0625	1	3.0625	1.26	0.2948
AB	7.5625	1	7.5625	3.10	0.1162
AC	0.0625	1	0.0625	0.03	0.8784
BC	1.5625	1	1.5625	0.64	0.4548
ABC	5.0625	1	5.0625	2.08	0.1875
Error	19.5000	8	2.4375		
Total	92.9375	15			

do el efecto de interacción es relativamente pequeño. El análisis de varianza, el cual se resume en la tabla 13-18, confirma la interpretación de las estimaciones de los efectos.

El programa Minitab hace el análisis de diseños factoriales 2^k. En la tabla 13-19 se muestra la salida del módulo DOE (*Design of Experiments*, Diseño de Experimentos) de Minitab para este experimento. En la parte superior de la tabla se presentan las estimaciones de los efectos y los coeficientes de regresión de cada efecto factorial. Sin embargo, se reporta un estadístico t para cada efecto en vez del estadístico F usado en la tabla 13-18. Ahora el cuadrado de la

Tabla 13-19 Análisis de Minitab del ejemplo 13-7

Efectos estimados y coeficientes de la rugosidad

Término	Efecto	Coef.	Coef. de desv. estándar	T	P
Constante		11.0625	0.3903	28.34	0.000
Alimentación	3.3750	1.6875	0.3903	4.32	0.003
Profundidad	1.6250	0.8125	0.3903	2.08	0.071
Ángulo	0.8750	0.4375	0.3903	1.12	0.295
Alimentación*Profundidad	1.3750	0.6875	0.3903	1.76	0.116
Alimentación*Ángulo	0.1250	0.0625	0.3903	0.16	0.877
Profundidad*Ángulo	–0.6250	–0.3125	0.3903	–0.80	0.446
Alimentación*Profundidad*Ángulo	1.1250	0.5625	0.3903	1.44	0.188

Análisis de varianza para la rugosidad

Fuente	GL	Sec SS	Aj SS	Aj MS	F	P
Efectos principales	3	59.188	59.188	19.729	8.09	0.008
Interacciones de dos factores	3	9.187	9.187	3.062	1.26	0.352
Interacciones de tres factores	1	5.062	5.062	5.062	2.08	0.188
Error residual	8	19.500	19.500	2.437		
Error puro	8	19.500	19.500	2.437		
Total	15	92.938				

variable aleatoria t con d grados de libertad es una variable aleatoria F con 1 y d grados de libertad en el numerador y el denominador, respectivamente. Por tanto, el cuadrado del estadístico t reportado por Minitab será igual (dejando de lado los errores de redondeo) al estadístico F de la tabla 13-18. Como ilustración, para el efecto principal de la alimentación, Minitab reporta $t = 4.32$ (con 8 grados de libertad), y $t^2 = (4.32)^2 = 18.66$, que es aproximadamente igual al cociente F para la alimentación reportada en la tabla 13-18 ($F = 18.69$). Este cociente F tiene 1 y 8 grados de libertad en el numerador y el denominador, respectivamente.

La sección inferior de la salida de Minitab de la tabla 13-19 es un resumen del análisis de varianza que se enfoca en los tipos de términos del modelo. Se usa el enfoque de un modelo de regresión en esta presentación. Quizá el lector encuentre útil repasar la sección 11-5.2, en particular el material sobre la prueba F parcial. El renglón rotulado "Efectos principales" que está abajo de "Fuente" se refiere a los tres efectos principales, alimentación, profundidad y ángulo, cada uno de ellos con un solo grado de libertad, dando el total 3 en la columna con encabezado "GL". La columna "Sec SS" (abreviatura de suma secuencial de cuadrados) reporta en qué cantidad se incrementa la suma de cuadrados del modelo cuando cada grupo de términos se agrega a un modelo que contiene los términos enlistados *arriba* de los grupos. El primer número de la columna "Sec SS" presenta la suma de cuadrados del modelo para ajustar un modelo que sólo incluye los tres efectos principales. El renglón rotulado "Interacciones de dos factores" se refiere a AB, AC y BC, y la suma secuencial de cuadrados reportada aquí es el incremento en la suma de cuadrados del modelo si los términos de las interacciones se agregan a un modelo que sólo contiene los efectos principales. De manera similar, la suma secuencial de cuadrados para la interacción de tres factores es el incremento en la suma de cuadrados del modelo que resulta de sumar el término ABC a un modelo que incluye todos los demás efectos. La columna con el encabezado "Aj SS" (abreviatura de suma ajustada de cuadrados) reporta la cantidad en que se incrementa la suma de cuadrados del modelo cuando cada grupo de términos se agrega a un modelo que incluye *todos* los demás términos. Entonces, puesto que cualquier diseño 2^k con el mismo número de réplicas en cada celda es un diseño ortogonal, resulta que la suma ajustada de cuadrados será igual a la suma secuencial de cuadrados. Por lo tanto, las pruebas F de cada renglón de la tabla del análisis de varianza de Minitab prueban la significación de cada grupo de términos (efectos principales, interacciones de dos factores, e interacciones de tres factores) como si fueran los últimos términos por incluir en el modelo. Evidentemente, sólo los términos de los efectos principales son significativos. Las pruebas t de los efectos individuales de los factores indican que la rapidez de alimentación y la profundidad del corte tienen efectos principales grandes, y quizás haya una ligera interacción entre estos dos factores. Por lo tanto, la salida de Minitab concuerda con los resultados presentados anteriormente.

Análisis residual

Los residuales de un diseño 2^k pueden obtenerse usando el método presentado aquí anteriormente para el diseño 2^2. Como un ejemplo, considérese el experimento de la rugosidad superficial. Los tres efectos más grandes son A, B y la interacción AB. El modelo de regresión usado para obtener los valores predichos es

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

donde x_1 representa el factor A, x_2 representa el factor B, y $x_1 x_2$ representa la interacción AB. Los coeficientes de regresión β_1, β_2 y β_{12} se estiman con la mitad de las estimaciones de los efectos correspondientes, y β_0 es el gran promedio. Por tanto,

$$\hat{y} = 11.0625 + \left(\frac{3.375}{2}\right)x_1 + \left(\frac{1.625}{2}\right)x_2 + \left(\frac{1.375}{2}\right)x_1x_2$$
$$= 11.0625 + 1.6875x_1 + 0.8125x_2 + 0.6875x_1x_2$$

Obsérvese que el programa Minitab presenta los coeficientes de regresión en la sección superior de la tabla 13-19. Los valores predichos se obtendrían sustituyendo los niveles bajo y alto de A y B en esta ecuación. Para ilustrar el punto, en la combinación de tratamientos donde A, B y C están todos en el nivel bajo, el valor predicho es

$$\hat{y} = 11.065 + 1.6875(-1) + 0.8125(-1) + 0.6875(-1)(-1)$$
$$= 9.25$$

Puesto que los valores observados en esta corrida son 9 y 7, los residuales son 9 – 9.25 = –0.25 y 7 – 9.25 = –2.25. Los residuales de las 14 corridas restantes se obtienen de manera similar.

En la figura 13-22 se muestra una gráfica de probabilidad normal de los residuales. Puesto que los residuales caen aproximadamente en una línea recta, no se sospecha ningún problema con la normalidad de los datos. No hay indicios de puntos atípicos serios. También sería conveniente graficar los residuales contra los valores predichos y contra cada uno de los factores A, B y C.

Otros métodos para juzgar la significación de los efectos

El análisis de varianza es una manera formal de determinar cuáles de los efectos son diferentes de cero. Otros dos métodos son útiles. En el primer método, pueden calcularse los errores estándar

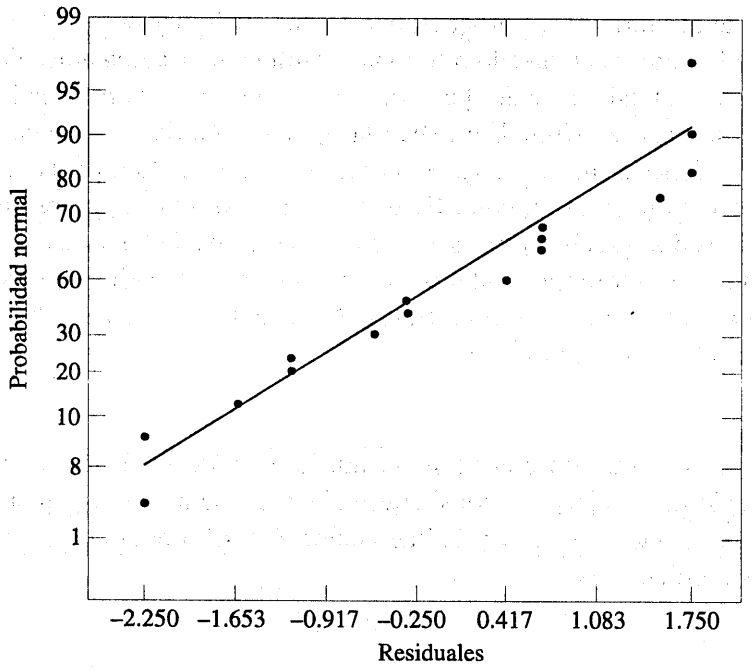


Figura 13-22 Gráfica de probabilidad normal de los residuales del experimento de la rugosidad superficial.

de los efectos y comparar la magnitud de los efectos con sus errores estándar. El segundo método utiliza gráficas de probabilidad normal para determinar la importancia de los efectos.

El error estándar de un efecto se encuentra con facilidad. Si se supone que hay n réplicas en cada una de las 2^k corridas del diseño, y si $y_{i1}, y_{i2}, \ldots, y_{in}$ son las observaciones de la corrida i-ésima, entonces

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i\cdot}\right)^2 \qquad i = 1, 2, \ldots, 2^k$$

es una estimación de la varianza en la corrida i-ésima. Las 2^k estimaciones de la varianza pueden agruparse para obtener una estimación de la varianza global

$$\hat{\sigma}^2 = \frac{1}{2^k(n-1)}\sum_{i=1}^{2^k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i\cdot}\right)^2 \tag{13-31}$$

Ésta es también la estimación de la varianza dada por el cuadrado medio de error en el procedimiento del análisis de varianza. La estimación de cada efecto tiene varianza dada por

$$\begin{aligned} V(\text{Efecto}) &= V\left[\frac{\text{Contraste}}{n2^{k-1}}\right] \\ &= \frac{1}{\left(n2^{k-1}\right)^2} V(\text{Contraste}) \end{aligned}$$

Cada contraste es una combinación lineal de los totales de 2^k tratamientos, y cada total consta de n observaciones. Por lo tanto,

$$V(\text{Contraste}) = n2^k\sigma^2$$

y la varianza de un efecto es

$$V(\text{Efecto}) = \frac{1}{\left(n2^{k-1}\right)^2} n2^k\sigma^2 = \frac{1}{n2^{k-2}}\sigma^2 \tag{13-32}$$

Por lo tanto, el error estándar de un efecto es

$$se(\text{Efecto}) = \sqrt{= \frac{1}{n2^{k-2}}\sigma^2}$$

El error estándar estimado de un efecto se encontraría sustituyendo σ^2 con su estimación, $\hat{\sigma}^2$, y sacando la raíz cuadrada de la ecuación 13-32.

Para ilustrar este enfoque para el experimento de la rugosidad superficial, se encuentra que $\hat{\sigma}^2 = 2.4375$, y el error estándar de cada efecto es

$$se(\text{Efecto}) = \sqrt{\frac{1}{n2^{k-2}}\sigma^2} = \sqrt{\frac{1}{2 \cdot 2^{3-2}}(2.4375)} = 0.78$$

Por lo tanto, dos límites del error estándar sobre las estimaciones de los efectos son

A:	3.375 ± 1.56	AC:	0.125 ± 1.56
B:	1.625 ± 1.56	BC:	-0.625 ± 1.56
C:	0.875 ± 1.56	ABC:	1.125 ± 1.56
AB:	1.375 ± 1.56		

Estos intervalos son intervalos de confianza de 95% aproximados. Indican que los dos efectos principales A y B son importantes, pero que los otros efectos no lo son, ya que los intervalos para todos los efectos excepto A y B incluyen al cero.

En la tabla 13-19, el programa Minitab reporta el error estándar de los coeficientes de regresión del modelo, $se(\hat{\beta})$. Entonces, puesto que la estimación del efecto es exactamente el doble del coeficiente de regresión del modelo,

$$se(\hat{\beta}) = \frac{1}{2} se(\text{Efecto})$$

Obsérvese que $se(\hat{\beta}) = 0.3903$ en la tabla 13-19. Por lo tanto, podrían construirse dos límites del error estándar sobre los coeficientes de regresión para evaluar la significación de los factores utilizados en el experimento.

Este enfoque de intervalos de confianza es una alternativa atractiva para el análisis de varianza usual. Con modificaciones relativamente simples, puede usarse en situaciones donde sólo se han hecho las réplicas de un número relativamente reducido de puntos del diseño. También pueden usarse gráficas de probabilidad normal para juzgar la significación de los efectos. Este método se ilustrará en la siguiente sección.

Proyección de diseños 2^k

Cualquier diseño 2^k se reducirá o proyectará en otro diseño 2^k con menos variables si uno o más de los factores originales se descartan. En ocasiones esto puede arrojar luz adicional sobre los factores que quedan. Por ejemplo, considérese el experimento de la rugosidad superficial. Puesto que el factor C y todas sus interacciones son insignificantes, el factor C podría eliminarse del diseño. El resultado es reducir el cubo de la figura 13-20 a un cuadrado en el plano A–B; por lo tanto, cada una de las cuatro corridas del nuevo diseño tiene cuatro réplicas. En general, si se eliminan h factores de tal suerte que queden $r = k - h$ factores, el diseño 2^k original con n réplicas se proyectará en un diseño 2^r con $n2^h$ réplicas.

13-6.3 Una sola réplica del diseño 2^k

Conforme aumenta el número de factores en un experimento factorial, el número de efectos que pueden estimarse también aumenta. Por ejemplo, un experimento 2^4 tiene 4 efectos principales, 6 interacciones de dos factores, 4 interacciones de tres factores y 1 interacción de cuatro factores, mientras que un experimento 2^6 tiene 6 efectos principales, 15 interacciones de dos factores, 20 interacciones de tres factores, 15 interacción de cuatro factores, 6 interacciones de cinco

factores y 1 interacción de seis factores. En la mayoría de las situaciones se aplica el **principio de efectos esparcidos**; es decir, el sistema suele estar dominado por los efectos principales y las interacciones de orden inferior. Por lo general las interacciones de tres factores y de órdenes superiores son insignificantes. Por lo tanto, cuando el número de factores es moderadamente grande, por ejemplo, $k \geq 4$ o 5, una práctica común es correr una sola réplica del diseño 2^k y agrupar o combinar después las interacciones de órdenes superiores como una estimación del error. En ocasiones a un diseño 2^k con una sola réplica se le llama diseño factorial 2^k **no replicado**.

Cuando se analizan datos de diseños factoriales no replicados, ocasionalmente ocurren interacciones reales de orden superior. El uso de un cuadrado medio de error que se obtiene agrupando las interacciones de órdenes superiores no es apropiado en estos casos. Es posible usar un método de análisis simple para salvar este problema. Se construye una gráfica de las estimaciones de los efectos en una escala de probabilidad normal. Los efectos que son insignificantes tienen una distribución normal con media cero y varianza σ^2, y tenderán a caer sobre una línea recta en esta gráfica, mientras que los efectos significativos tendrán medias diferentes de cero y no insignificantes sobre la línea recta. Este método se ilustrará en el siguiente ejemplo.

EJEMPLO 13-8

En un artículo de *Solid State Technology* ("Diseño ortogonal para optimización de procesos y su aplicación en el grabado químico con plasma", pp. 127-132) se describe la aplicación de diseños factoriales en el desarrollo de un proceso de grabado químico para nitruros en un dispositivo de grabado químico con plasma para una sola oblea. El proceso usa C_2F_6 como gas de reacción. Es posible variar el flujo de gas, la energía aplicada al cátodo, la presión en la cámara del reactor, y la separación entre el ánodo y el cátodo (entrehierro). En general, serían de interés varias variables de respuesta en este proceso, pero en este ejemplo nos enfocaremos en la rapidez de grabado para el nitruro de silicio.

Se usará una sola réplica de un diseño 2^4 para investigar este proceso. Puesto que es improbable que las interacciones de tres y cuatro factores sean significativas, tentativamente se planea combinarlas como una estimación del error. Los niveles de los factores usados en el diseño se muestran a continuación:

	Factor del diseño			
Nivel	Entrehierro (cm)	Presión (mTorr)	Flujo de C_2F_6 (SCCM)	Energía (w)
Bajo (–)	0.80	450	125	275
Alto (+)	1.20	550	200	325

En la tabla 13-20 se presentan los datos de las 16 corridas del diseño 2^4. La tabla 13-21 es la tabla de los signos positivos y negativos del diseño 2^4. Los signos de las columnas de esta tabla pueden usarse para estimar los efectos de los factores. Por ejemplo, la estimación del factor A es

$$\begin{aligned} A &= \frac{1}{8}[a + ab + ac + abc + ad + abd + acd + abcd - (1) - b \\ &\quad - c - bc - d - bd - cd - bcd] \\ &= \frac{1}{8}[669 + 650 + 642 + 635 + 749 + 868 + 860 + 729 \\ &\quad - 550 - 604 - 633 - 601 - 1\,037 - 1\,052 - 1\,075 - 1\,063] \\ &= -101.625 \end{aligned}$$

Tabla 13-20 El diseño 2^4 del experimento del grabado químico con plasma

A (Entrehierro)	*B* (Presión)	*C* (Flujo de C_2F_6)	*D* (Energía)	Rapidez del grab. (Å/min)
−1	−1	−1	−1	550
1	−1	−1	−1	669
−1	1	−1	−1	604
1	1	−1	−1	650
−1	−1	1	−1	633
1	−1	1	−1	642
−1	1	1	−1	601
1	1	1	−1	635
−1	−1	−1	1	1 037
1	−1	−1	1	749
−1	1	−1	1	1052
1	1	−1	1	868
−1	−1	1	1	1 075
1	−1	1	1	860
−1	1	1	1	1 063
1	1	1	1	729

Tabla 13-21 Constantes de contraste del diseño 2^4

	A	*B*	*AB*	*C*	*AC*	*BC*	*ABC*	*D*	*AD*	*BD*	*ABD*	*CD*	*ACD*	*BCD*	*ABCD*
(1)	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+
a	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−
b	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−
ab	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+
c	−	−	+	+	−	−	+	−	+	+	−	−	+	+	−
ac	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+
bc	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+
abc	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−
d	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−
ad	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+
bd	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+
abd	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−
cd	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+
acd	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−
bcd	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
abcd	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

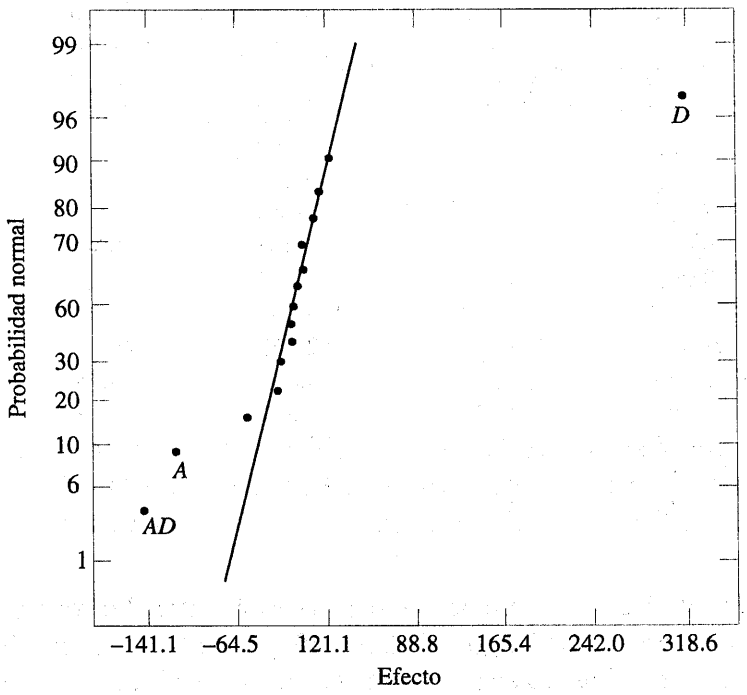


Figura 13-23 Gráfica de probabilidad normal de los efectos del experimento del grabado químico con plasma.

Por tanto, el efecto de aumentar la separación entre el ánodo y el cátodo de 0.80 cm a 1.20 cm es disminuir la rapidez de grabado en 101.625 Å/min.

Es sencillo verificar (utilizando Minitab, por ejemplo) que el conjunto completo de las estimaciones de los efectos es

A	=	−101.625	AD	=	−153.625
B	=	−1.625	BD	=	−0.625
AB	=	−7.875	ABD	=	4.125
C	=	7.375	CD	=	−2.125
AC	=	−24.875	ACD	=	5.625
BC	=	−43.875	BCD	=	−25.375
ABC	=	−15.625	$ABCD$	=	−40.125
D	=	306.125			

En la figura 13-23 se muestra la gráfica de probabilidad normal de estos efectos del experimento del grabado químico con plasma. Evidentemente, los efectos principales de A y D y la interacción AD son significativos, pues se apartan bastante de la línea que pasa por los demás puntos. El análisis de varianza que se resume en la tabla 13-22 confirma estos hallazgos. Obsérvese que en el análisis de varianza se han agrupado las interacciones de tres y cuatro factores para formar el cuadrado medio de error. Si la gráfica de probabilidad normal hubiera indicado que alguna de estas interacciones era importante, entonces no se habría incluido en el término del error.

Tabla 13-22 Análisis de varianza del experimento del grabado químico con plasma

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
A	41 310.563	1	41 310.563	20.28	0.0064
B	10.563	1	10.563	<1	—
C	217.563	1	217.563	<1	—
D	374 850.063	1	374 850.063	183.99	0.0000
AB	248.063	1	248.063	<1	—
AC	2 475.063	1	2 475.063	1.21	0.3206
AD	94 402.563	1	94 402.563	46.34	0.0010
BC	7 700.063	1	7 700.063	3.78	0.1095
BD	1.563	1	1.563	<1	—
CD	18.063	1	18.063	<1	—
Error	10 186.813	5	2 037.363		
Total	531 420.938	15			

Puesto que $A = -101.625$, el efecto de incrementar la separación entre el cátodo y el ánodo es disminuir la rapidez de grabado. Sin embargo, $D = 306.125$; por tanto, la aplicación de niveles más altos de energía incrementará la rapidez de grabado. La figura 13-24 es una gráfica de la interacción *AD*. Esta gráfica indica que es reducido el efecto de cambiar la anchura del entrehierro con ajustes de energía bajos, pero que incrementar el entrehierro con ajustes de energía altos produce una reducción sustancial de la rapidez de grabado. Los valores de rapidez de grabado altos se obtienen con ajustes de energía altos y anchuras estrechas del entrehierro.

Los residuales del experimento pueden obtenerse a partir del modelo de regresión

$$\hat{y} = 776.0625 - \left(\frac{101.625}{2}\right)x_1 + \left(\frac{306.125}{2}\right)x_4 - \left(\frac{153.625}{2}\right)x_1x_4$$

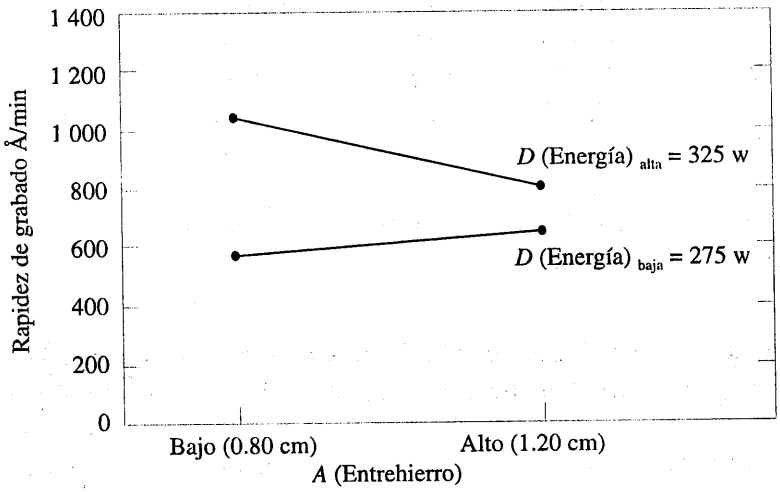


Figura 13-24 Interacción *AD* (Entrehierro-Energía) del experimento del grabado químico con plasma.

Por ejemplo, cuando tanto A como D están en el nivel bajo, el valor predicho es

$$\hat{y} = 776.0625 - \left(\frac{101.625}{2}\right)(-1) + \left(\frac{306.125}{2}\right)(-1) - \left(\frac{153.625}{2}\right)(-1)(-1)$$
$$= 597$$

y los cuatro residuales de esta combinación de tratamientos son

$$e_1 = 550 - 597 = -47$$
$$e_2 = 604 - 597 = 7$$
$$e_3 = 633 - 597 = 36$$
$$e_4 = 601 - 597 = 4$$

Los residuales de las otras tres combinaciones de tratamientos (A alta, D baja), (A baja, D alta) y (A alta, D alta), se obtienen de manera similar. En la figura 13-25 se muestra una gráfica de probabilidad normal de los residuales. La gráfica es satisfactoria.

13-6.4 Adición de los puntos centrales a un diseño 2^k

Una preocupación potencial en el uso de diseños factoriales con dos niveles es el supuesto de linealidad de los efectos de los factores. Desde luego, no se requiere una linealidad perfecta, y el

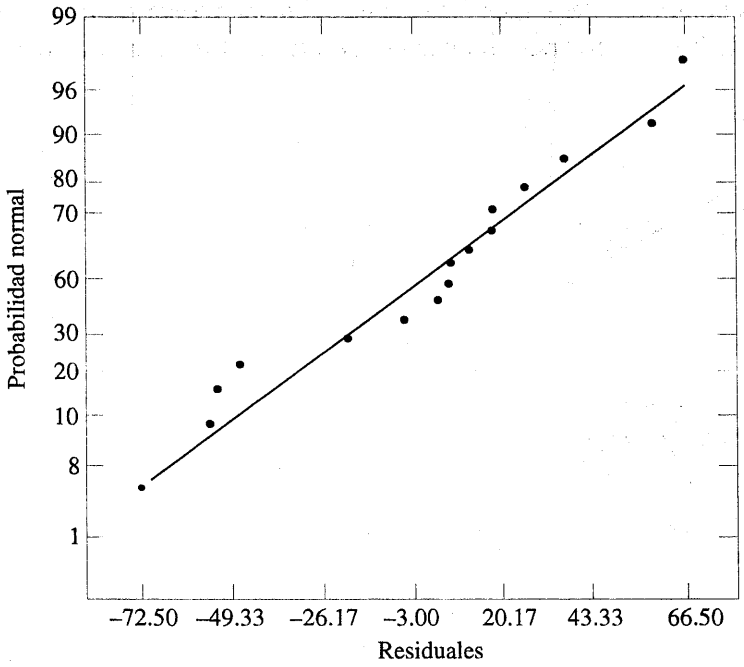


Figura 13-25 Gráfica de probabilidad normal de los residuales del experimento del grabado químico con plasma.

sistema 2^k funcionará bastante bien incluso cuando la validez del supuesto de linealidad sea tan sólo aproximada. Sin embargo, existe un método para hacer réplicas en ciertos puntos del factorial 2^k que proporcionarán protección contra la curvatura a la vez que permitirán una estimación independiente del error que deberá obtenerse. El método consiste en agregar **puntos centrales** al diseño 2^k. Éstos consisten en n_C réplicas corridas en el punto $x_i = 0$ ($i = 1, 2, ..., k$). Una razón importante para agregar las corridas de réplicas en el centro del diseño es que los puntos centrales no afectan las estimaciones convencionales de los efectos de un diseño 2^k. Se supone que los k factores son cuantitativos.

Para ilustrar el enfoque, considérese un diseño 2^2 con una observación en cada uno de los puntos factoriales (–, –), (+, –), (–, +) y (+, +), y n_C observaciones en los puntos centrales (0, 0). En la figura 13-26 se ilustra la situación. Sea $\bar{y}_F$ el promedio de las cuatro corridas en los cuatro puntos factoriales, y sea $\bar{y}_C$ el promedio de las n_C corridas en el punto central. Si la diferencia $\bar{y}_F - \bar{y}_C$ es pequeña, entonces los puntos centrales caen en el plano que pasa por los puntos factoriales, o cerca del mismo, y no hay curvatura. Por otra parte, si $\bar{y}_F - \bar{y}_C$ es grande, entonces hay curvatura. Una suma de cuadrados con un solo grado de libertad para la curvatura está dada por

$$SS_{\text{Curvatura}} = \frac{n_F n_C \left(\bar{y}_F - \bar{y}_C\right)^2}{n_F + n_C} = \left(\frac{\bar{y}_F - \bar{y}_C}{\sqrt{\dfrac{1}{n_F} + \dfrac{2}{n_C}}} \right)^2 \tag{13-33}$$

donde, en general, n_F es el número de puntos factoriales del diseño. Esta cantidad puede compararse con el cuadrado medio de error para probar la curvatura. Obsérvese que cuando la ecua-

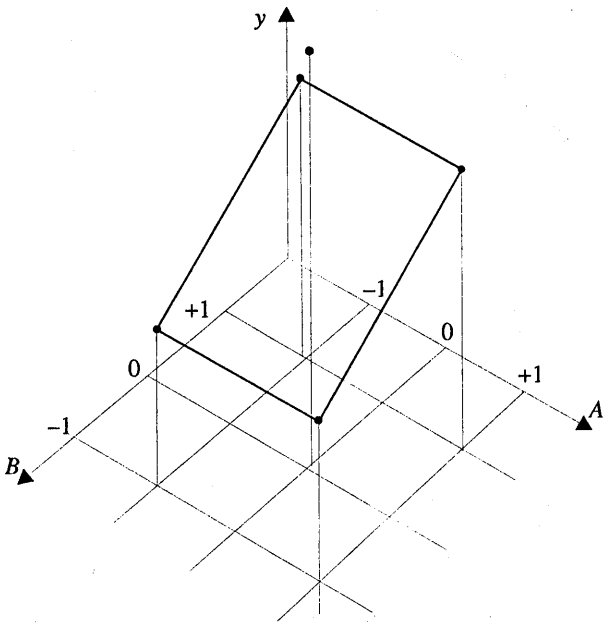


Figura 13-26 Diseño 2^2 con puntos centrales.

ción 13-33 se divide por $\hat{\sigma}^2 = MS_E$ el resultado es similar al cuadrado del estadístico t utilizado para comparar dos medias.

Más específicamente, cuando se agregan puntos en el centro del diseño 2^k, el modelo que puede tomarse en consideración es

$$Y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum\sum_{i<j} \beta_{ij} x_i x_j + \sum_{j=1}^{k} \beta_{jj} x_j^2 + \varepsilon$$

donde las β_{jj} son efectos cuadráticos puros. En la prueba de la curvatura en realidad se prueban las hipótesis

$$H_0 : \sum_{j=1}^{k} \beta_{jj} = 0$$

$$H_1 : \sum_{j=1}^{k} \beta_{jj} \neq 0$$

Además, si los puntos factoriales del diseño son no replicados, pueden usarse n_C puntos centrales para construir una estimación del error con $n_C - 1$ grados de libertad.

EJEMPLO 13-9

Una ingeniera química estudia la conversión porcentual o rendimiento de un proceso. Hay dos variables de interés, el tiempo de reacción y la temperatura de reacción. Debido a que no está segura acerca del supuesto de linealidad en la región de exploración, la ingeniera decide conducir un diseño 2^2 (con una sola réplica de cada corrida factorial) con la adición de cinco puntos centrales. El diseño y los datos del rendimiento se presentan en la figura 13-27.

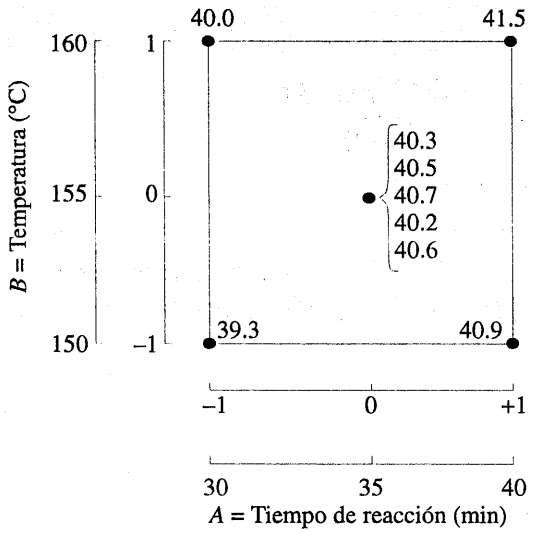


Figura 13-27 El diseño 2^2 con cinco puntos centrales del ejemplo 13-9.

Tabla 13-23 Análisis de varianza del ejemplo 13-9

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
A (Tiempo)	2.4025	1	2.4025	55.87	0.0017
B (Temperatura)	0.4225	1	0.4225	9.83	0.0350
AB	0.0025	1	0.0025	0.06	0.8237
Curvatura	0.0027	1	0.0027	0.06	0.8163
Error	0.1720	4	0.0430		
Total	3.0022	8			

En la tabla 13-23 se resume el análisis de varianza de este experimento. El cuadrado medio de error se calcula a partir de los puntos centrales de la siguiente manera:

$$MS_E = \frac{SS_E}{n_C - 1} = \frac{\sum_{\text{Puntos centrales}} (y_i - \bar{y}_C)^2}{n_C - 1}$$

$$= \frac{\sum_{i=1}^{5} (y_i - 40.46)^2}{4} = \frac{0.1720}{4} = 0.0430$$

El promedio de los puntos de la porción factorial del diseño es $\bar{y}_F = 40.425$, y el promedio de los puntos del centro es $\bar{y}_C = 40.46$. La diferencia $\bar{y}_F - \bar{y}_C = 40.425 - 40.46 = -0.035$ parece ser pequeña. La suma de los cuadrados de la curvatura de la tabla del análisis de varianza se calcula con la ecuación 13-33 como sigue:

$$SS_{\text{Curvatura}} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C}$$

$$= \frac{(4)(5)(-0.035)^2}{4+5}$$

$$= 0.0027$$

El análisis de varianza indica que ambos factores presentan efectos principales significativos, que no hay ninguna interacción y que no hay evidencia de curvatura de la respuesta en la región de exploración. Es decir, la hipótesis nula H_0: $\sum_{j=1}^{k} \beta_{jj} = 0$ no puede rechazarse.

EJERCICIOS DE LA SECCIÓN 13-6

13-13. Un ingeniero se interesa en el efecto de la rapidez de corte (A), la dureza del metal (B) y el ángulo de corte (C) sobre la vida de una herramienta de corte. Se escogen dos niveles de cada factor, y se realizan dos réplicas de un diseño factorial 2^3. En la tabla siguiente se presentan los datos de la vida de la herramienta de corte (en horas).

Combinación de tratamientos	Réplicas I	Réplicas II
(1)	221	311
a	325	435
b	354	348
ab	552	472
c	440	453
ac	406	377
bc	605	500
abc	392	419

a) Analice los datos de este experimento.
b) Encuentre un modelo de regresión apropiado que explique la vida de la herramienta en términos de las variables usadas en el experimento.
c) Analice los residuales de este experimento.

13-14. Se piensa que cuatro factores influyen en el sabor de una bebida refrescante: el tipo de edulcorante (*A*), la proporción del jarabe con el agua (*B*), el nivel de carbonación (*C*), y la temperatura (*D*). Cada factor puede ajustarse en dos niveles, produciendo un diseño 2^4. En cada corrida del diseño, se hacen llegar muestras de la bebida a un panel de prueba compuesto por 20 personas. Cada integrante asigna a la bebida una calificación en puntos del 1 al 10. La puntuación total es la variable de respuesta, y el objetivo es encontrar la formulación que maximice la puntuación total. Se corren dos réplicas de este diseño, y los resultados se muestran a continuación. Analice los datos y sacar conclusiones. Use $\alpha = 0.05$ en las pruebas estadísticas.

Combinación de tratamientos	Réplicas I	Réplicas II
(1)	159	163
a	168	175
b	158	163
ab	166	168
c	175	178
ac	179	183
bc	173	168
abc	179	182
d	164	159
ad	187	189
bd	163	159
abd	185	191
cd	168	174
acd	197	199
bcd	170	174
abcd	194	198

13-15. Considérese el experimento del ejercicio 13-14. Determine un modelo apropiado y grafique los residuales contra los niveles de los factores *A*, *B*, *C* y *D*. Construya también una gráfica de probabilidad normal de los residuales. Comente estas gráficas.

13-16. Encuentre el error estándar de los efectos del experimento del ejercicio 13-14. Utilizando los errores estándar como guía, ¿qué factores parecen ser significativos?

13-17. Los datos que se muestran aquí representan una sola réplica de un diseño 2^5 que se usa en un experimento para estudiar la resistencia a la compresión del concreto. Los factores son la mezcla (*A*), el tiempo (*B*), el laboratorio (*C*), la temperatura (*D*) y el tiempo de fraguado (*E*).

(1)	= 700	*e*	= 800
a	= 900	*ae*	= 1 200
b	= 3 400	*be*	= 3 500
ab	= 5 500	*abe*	= 6 200
c	= 600	*ce*	= 600
ac	= 1 000	*ace*	= 1 200
bc	= 3 000	*bce*	= 3 006
abc	= 5 300	*abce*	= 5 500
d	= 1 000	*de*	= 1 900
ad	= 1 100	*ade*	= 1 500
bd	= 3 000	*bde*	= 4 000
abd	= 6 100	*abde*	= 6 500
cd	= 800	*cde*	= 1 500
acd	= 1 100	*acde*	= 2 000
bcd	= 3 300	*bcde*	= 3 400
abcd	= 6 000	*abcde*	= 6 800

a) Estime los efectos de los factores.
b) ¿Qué efectos parecen ser importantes?

Use una gráfica de probabilidad normal.

c) Si es deseable maximizar la resistencia a la compresión, ¿en qué dirección se ajustarían las variables del proceso?

d) Analice los residuales de este experimento.

13-18. En un artículo de *IEEE Transactions on Semiconductor Manufacturing* (Vol. 5, No. 3, pp. 214-222) se describe un experimento para investigar la carga superficial sobre una oblea de silicio. Se piensa que los factores que influyen en la carga superficial inducida son el método de limpieza (secado centrifugado con enjuague o SCE y secado centrifugado o SC) y la posición de la oblea donde se midió la carga. Los datos de respuesta de la carga superficial ($\times 10^{11}$ q/cm^3) se muestran abajo.

		Posición de prueba	
		Izquierda	**Derecha**
Método de limpieza	**SC**	1.66	1.84
		1.90	1.84
		1.92	1.62
	SCE	–4.21	–7.58
		–1.35	–2.20
		–2.08	–5.36

a) Estime los efectos de los factores.

b) ¿Qué factores parecen ser importantes? Use $\alpha = 0.05$.

c) Analice los residuales de este experimento.

13-19. Un experimento descrito por M.G. Natrella en *Handbook of Experimental Statistics* del National Bureau of Standards (No. 91) se ocupa de la prueba de resistencia a la flama para las telas después de aplicar tratamientos para retardar el fuego. Los cuatro factores considerados son el tipo de tela (A), el tipo de tratamiento para retardar el fuego (B), la condición de lavado (C, el nivel bajo es sin lavado, el nivel alto es después de una lavada) y el método de flama empleado la prueba de (D). Todos los factores corren en dos niveles, y la variable de respuesta son las pulgadas de tela quemadas en una muestra de tamaño estándar. Los datos son:

(1)	=	42	d	=	40
a	=	31	ad	=	30
b	=	45	bd	=	50
ab	=	29	abd	=	25
c	=	39	cd	=	40
ac	=	28	acd	=	25
bc	=	46	bcd	=	50
abc	=	32	$abcd$	=	23

a) Estime los efectos y trazar una gráfica normal de los efectos.

b) Construya una gráfica de probabilidad normal de los residuales y comente los resultados.

c) Construya una tabla del análisis de varianza suponiendo que las interacciones de tres y cuatro factores son insignificantes.

13-20. Se corrió un experimento en una planta de fabricación de semiconductores en un esfuerzo por incrementar el rendimiento. Se estudiaron cinco factores, cada uno con dos niveles. Los factores (y los niveles) fueron A = abertura del orificio (pequeño, grande), B = tiempo de exposición (20% abajo del nominal, 20% arriba del nominal), C = tiempo de desarrollo (30 s, 45 s), D = dimensiones de la plantilla (pequeña, grande), y E = tiempo de grabado químico (14.5 min, 15.5 min). Se corrió el diseño 2^5 no replicado que se muestra abajo.

(1)	=	7	e	=	8
a	=	9	ae	=	12
b	=	34	be	=	35
ab	=	55	abe	=	52
c	=	16	ce	=	15
ac	=	20	ace	=	22
bc	=	40	bce	=	45
abc	=	60	$abce$	=	65
d	=	8	de	=	6
ad	=	10	ade	=	10
bd	=	32	bde	=	30
abd	=	50	$abde$	=	53
cd	=	18	cde	=	15
acd	=	21	$acde$	=	20
bcd	=	44	$bcde$	=	41
$abcd$	=	61	$abcde$	=	63

a) Construya una gráfica de probabilidad normal de las estimaciones de los efectos. ¿Qué efectos parecen ser grandes?
b) Conduzca un análisis de varianza para confirmar los hallazgos del inciso a).
c) Grafique los residuales en papel probabilidad normal. ¿La gráfica es satisfactoria?
d) Grafique los residuales contra los rendimientos predichos y contra cada uno de los cinco factores. Comente las gráficas.
e) Interprete cualquier interacción significativa.
f) ¿Cuáles serían las recomendaciones respecto de las condiciones de operación del proceso?
g) Proyecte el diseño 2^5 de este problema en un diseño 2^r para $r < 5$ con los factores importantes. Grafique el diseño y muestre el promedio y el rango de los rendimientos de cada corrida. ¿Esta gráfica es de ayuda para interpretar los datos?

13-21. Considérese el experimento de los semiconductores del ejercicio 13-20. Suponga que fue posible agregar un punto central (con cinco réplicas) a este diseño y que las respuestas en el centro son 45, 40, 41, 47 y 43.
a) Pruebe la falta de ajuste de este modelo, usando $\alpha = 0.05$.
b) Estime el error experimental usando los puntos centrales. ¿Cómo se compara esta estimación con la estimación obtenida en el ejercicio 13-20?

13-22. Considérense los datos del ejercicio 13-13, únicamente los de la repetición I. Suponga que un punto central (con cuatro réplicas) se agrega a estas ocho corridas. La respuesta de la vida de la herramienta en el punto central es 425, 400, 437 y 418.
a) Estime los efectos de los factores.
b) Pruebe la falta de ajuste, usando $\alpha = 0.05$.
c) Pruebe los efectos principales y los efectos de las interacciones, usando $\alpha = 0.05$.
d) Analice los residuales de este experimento.

13-7 FORMACIÓN DE BLOQUES Y CONFUSIÓN EN EL DISEÑO 2^k

Con frecuencia es imposible hacer todas las observaciones de un diseño factorial 2^k bajo condiciones homogéneas. La formación de bloques es la técnica de diseño que resulta apropiada para esta situación general. Sin embargo, en muchas situaciones el tamaño de los bloques es menor que el número de corridas en una réplica completa. En estos casos, la **confusión** o **mezclado** es un procedimiento conveniente para correr el diseño 2^k en 2^p bloques donde el número de corridas en un bloque es menor que el número de combinaciones de tratamientos de una réplica completa. La técnica hace que ciertos efectos de interacción no se distingan de los bloques, es decir, hace que **se confundan (o se mezclen) con los bloques**. La confusión se ilustrará en el diseño factorial 2^k en 2^p bloques, donde $p < k$.

Considérese un diseño 2^2. Suponga que cada una de las $2^2 = 4$ combinaciones de tratamientos requiere cuatro horas de análisis de laboratorio. Por tanto, se necesitarán dos días para llevar a cabo el experimento. Si los días se consideran como bloques, entonces deberán asignarse dos de las cuatro combinaciones de tratamientos a cada día.

Este diseño se ilustra en la figura 13-28. Obsérvese que el bloque 1 contiene las combinaciones de tratamientos (1) y ab y que el bloque 2 contiene a y b. Los contrastes para estimar los efectos principales de los factores A y B son

$$\text{Contraste}_A = ab + a - b - (1)$$
$$\text{Contraste}_B = ab + b - a - (1)$$

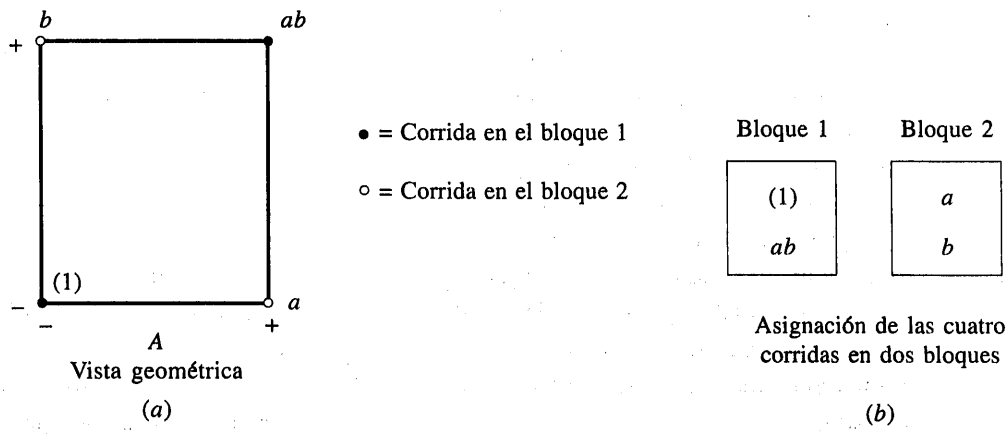


Figura 13-28 Un diseño 2^2 en dos bloques. ***a***) Vista geométrica. ***b***) Asignación de las cuatro corridas en dos bloques.

Obsérvese que estos contrastes no son afectados por la separación en bloques ya que en cada contraste hay una combinación de tratamientos positivos y negativos de cada bloque. Es decir, cualquier diferencia entre el bloque 1 y el bloque 2 que incremente las lecturas en un bloque por una constante aditiva se cancela. El contraste de la interacción AB es

$$\text{Contraste}_{AB} = ab + (1) - a - b$$

Puesto que las dos combinaciones de tratamientos con los signos positivos, ab y (1), están en el bloque 1 y las dos con los signos negativos, a y b, están en el bloque 2, el efecto del bloque y la interacción AB son idénticos. Es decir, la interacción AB está confundida con los bloques.

La razón de esto se pone de manifiesto en la tabla de signos positivos y negativos del diseño 2^2 que se muestra en la tabla 13-13. En la tabla se ve que todas las combinaciones de tratamientos que tienen un signo positivo para AB se asignan al bloque 1, mientras que todas las combinaciones de tratamientos que tienen un signo negativo para AB se asignan al bloque 2.

Este esquema puede usarse para confundir cualquier diseño 2^k en dos bloques. Como un segundo ejemplo, considérese un diseño 2^3 que se corre en dos bloques. Por la tabla de signos positivos y negativos, presentada en la tabla 13-16, las combinaciones de tratamientos que tienen un signo negativo en la columna ABC se asignan al bloque 1 y las que tienen un signo positivo en esa columna se asignan al bloque 2. El diseño resultante se muestra en la figura 13-29.

Hay un método más general para construir los bloques. El método emplea un **contraste de definición**, es decir,

$$L = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k \tag{13-34}$$

donde x_i es el nivel del factor i-ésimo que figura en una combinación de tratamientos y α_i es el exponente que figura en el factor i-ésimo en el efecto que va a confundirse con los bloques. Para el sistema 2^k, se tiene $\alpha_i = 0$ o 1, y $x_i = 0$ (nivel bajo) o $x_i = 1$ (nivel alto). Las combinaciones de tratamientos que producen el mismo valor de L (módulo 2) se colocarán en el mismo bloque.

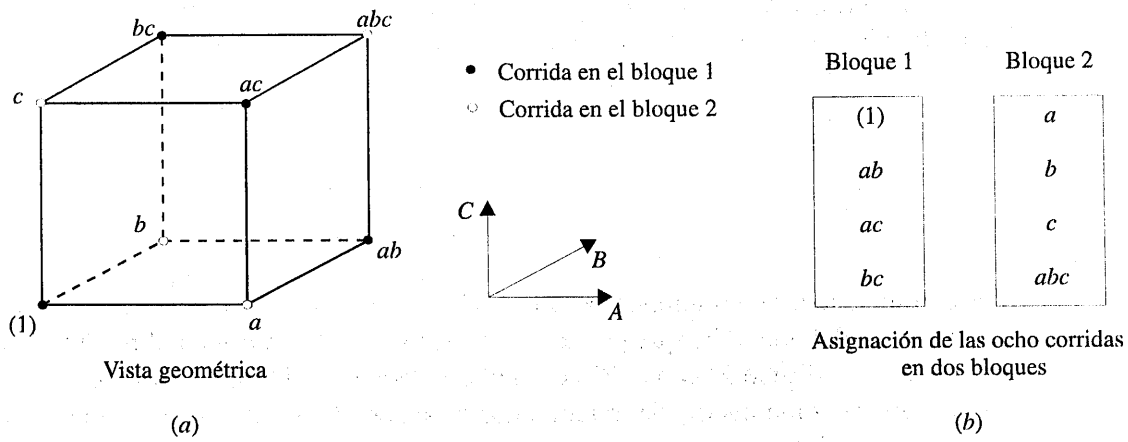


Figura 13-29 El diseño 2^3 en dos bloques con ABC confundidos. ***a***) Vista geométrica. ***b***) Asignación de las ocho corridas en dos bloques.

Puesto que los únicos valores posibles de L (mod 2) son 0 y 1, con esto se asignarán las 2^k combinaciones de tratamientos en dos bloques exactamente.

Como un ejemplo, considérese el diseño 2^3 con ABC confundido con los bloques. Aquí x_1 corresponde a A, x_2 a B, x_3 a C, y $\alpha_1 = \alpha_2 = \alpha_3 = 1$. Por tanto, el contraste de definición que se usaría para confundir ABC con los bloques es

$$L = x_1 + x_2 + x_3$$

Para asignar las combinaciones de tratamientos a los dos bloques, se sustituyen las combinaciones de tratamientos en el contraste de definición de la siguiente manera:

$$\begin{aligned}
(1):&\quad L = 1(0) + 1(0) + 1(0) = 0 = 0 \text{ (modelo 2)}\\
a:&\quad L = 1(1) + 1(0) + 1(0) = 1 = 1 \text{ (modelo 2)}\\
b:&\quad L = 1(0) + 1(1) + 1(0) = 1 = 1 \text{ (modelo 2)}\\
ab:&\quad L = 1(1) + 1(1) + 1(0) = 2 = 0 \text{ (modelo 2)}\\
c:&\quad L = 1(0) + 1(0) + 1(1) = 1 = 1 \text{ (modelo 2)}\\
ac:&\quad L = 1(1) + 1(0) + 1(1) = 2 = 0 \text{ (modelo 2)}\\
bc:&\quad L = 1(0) + 1(1) + 1(1) = 2 = 0 \text{ (modelo 2)}\\
abc:&\quad L = 1(1) + 1(1) + 1(1) = 3 = 1 \text{ (modelo 2)}
\end{aligned}$$

Por tanto (1), ab, ac y bc se corren en el bloque 1, y a, b y c y abc se corren en el bloque 2. Este mismo diseño se ilustra en la figura 13-29.

Para construir estos diseños, resulta conveniente un método abreviado. Al bloque que contiene la combinación de tratamientos (1) se le llama el **bloque principal**. Cualquier elemento [con excepción de (1)] en el bloque principal puede generarse multiplicando otros dos elemen-

tos del bloque principal módulo 2 para los exponentes. Por ejemplo, considérese el bloque principal del diseño 2^3 con ABC confundido, el cual se ilustra en la figura 13-29. Obsérvese que

$$ab \cdot ac = a^2bc = bc$$
$$ab \cdot bc = ab^2c = ac$$
$$ac \cdot bc = abc^2 = ab$$

Las combinaciones de tratamientos del otro bloque (o bloques) pueden generarse multiplicando un elemento del nuevo bloque por cada elemento del bloque principal módulo 2 para los exponentes. Para el diseño 2^3 con ABC confundido, dado que el bloque principal es (1), ab, ac y bc, se sabe que la combinación de tratamientos b está en el otro bloque. Por tanto, los elementos de este segundo bloque son

$$\begin{aligned} b \cdot (1) &= b \\ b \cdot ab &= ab^2 = a \\ b \cdot ac &= abc \\ b \cdot bc &= b^2c = c \end{aligned}$$

EJEMPLO 13-10

Se conduce un experimento para investigar el efecto de cuatro factores sobre la pérdida de alcance de un misil tierra-aire para cargar al hombro. Los cuatro factores son el tipo de objetivo (A), el tipo de buscador (B), la altitud del objetivo (C) y el rango del objetivo (D). Cada factor puede correrse convenientemente en dos niveles, y el sistema de rastreo óptico permite medir la pérdida de alcance con aproximación de pies. Se usan dos operadores o artilleros diferentes en el vuelo de prueba y, puesto que puede haber diferencias entre los operadores, los ingenieros de la prueba decidieron llevar a cabo el diseño 2^4 en dos bloques con $ABCD$ confundido. Por tanto, el contraste de definición es

$$L = x_1 + x_2 + x_3 + x_4$$

En la figura 13-30 se muestra el diseño experimental y los datos resultantes. Las estimaciones de los efectos obtenidas con el programa Minitab se presentan en la tabla 13-24. La gráfica de probabilidad normal de los efectos en la figura 13-31 revela que A (el tipo de objetivo), D (el rango del objetivo), AD y AC tienen los efectos más grandes. En la tabla 13-25 se muestra un análisis de varianza de confirmación, en el cual las interacciones de tres factores se agrupan como error. Puesto que las interacciones AC y AD son significativas, es razonable concluir que A (el tipo de objetivo), C (la altitud del objetivo) y D (el rango del objetivo) tienen efectos importantes sobre la pérdida en el alcance y que hay interacciones entre el tipo y la altitud del objetivo y entre el tipo y el rango del objetivo. Obsérvese que el efecto $ABCD$ se trata como bloques en este análisis.

Es posible confundir el diseño 2^k en cuatro bloques con 2^{k-2} observaciones cada uno. Para construir el diseño, se eligen dos efectos para confundirlos con los bloques, y se obtienen sus

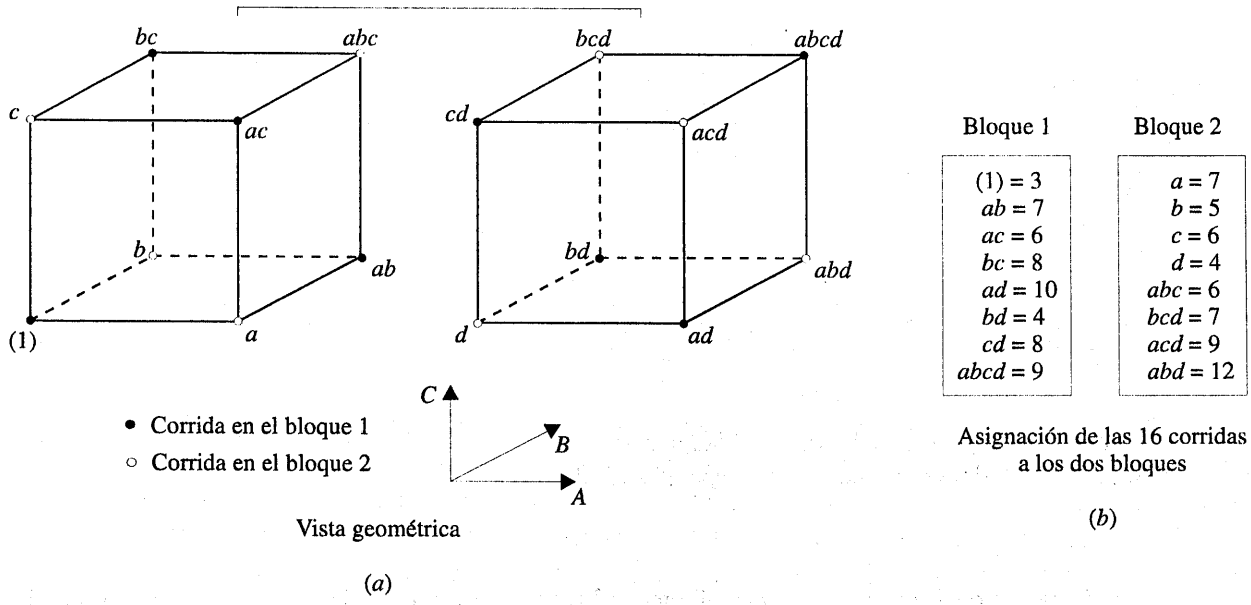


Figura 13-30 El diseño 2^4 en dos bloques del ejemplo 13-10. ***a***) Vista geométrica. ***b***) Asignación de las 16 corridas a los dos bloques.

Tabla 13-24 Estimaciones de los efectos del programa Minitab para el ejemplo 13-10

Efectos estimados y coeficientes del alcance		
Término	Efecto	Coeficiente
Constante		6.938
Bloque		0.063
A	2.625	1.312
B	0.625	0.313
C	0.875	0.438
D	1.875	0.938
A*B	–0.125	–0.063
A*C	–2.375	–1.187
A*D	1.625	0.813
B*C	–0.375	–0.188
B*D	–0.375	–0.187
C*D	–0.125	–0.062
A*B*C	–0.125	–0.063
A*B*D	0.875	0.438
A*C*D	–0.375	–0.187
B*C*D	–0.375	–0.187

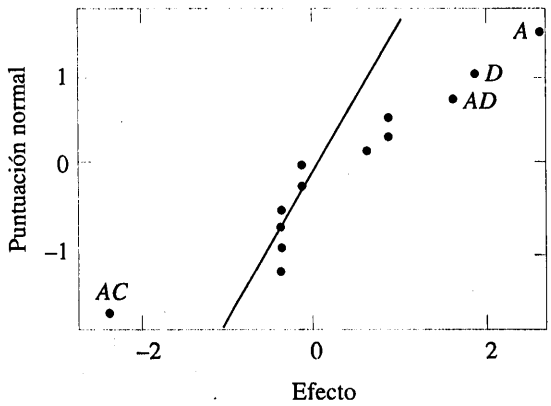


Figura 13-31 Gráfica de probabilidad normal de los efectos del programa Minitab, ejemplo 13-10.

contrastes de definición. Un tercer efecto, la **interacción generalizada** de los dos efectos elegidos inicialmente, también se confunde con los bloques. La interacción generalizada de dos efectos se encuentra multiplicando sus respectivas letras y reduciendo los exponentes módulo 2.

Por ejemplo, considérese el diseño 2^4 en cuatro bloques. Si AC y BD se confunden con los bloques, su interacción generalizada es $(AC)(BD) = ABCD$. El diseño se construye utilizando los contrastes de definición para AC y BD:

$$L_1 = x_1 + x_3$$
$$L_2 = x_2 + x_4$$

Tabla 13-25 Análisis de varianza del ejemplo 13-10

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Bloques ($ABCD$)	0.0625	1	0.0625	0.06	—
A	27.5625	1	27.5625	25.94	0.0070
B	1.5625	1	1.5625	1.47	0.2920
C	3.0625	1	3.0625	2.88	0.1648
D	14.0625	1	14.0625	13.24	0.0220
AB	0.0625	1	0.0625	0.06	—
AC	22.5625	1	22.5625	21.24	0.0100
AD	10.5625	1	10.5625	9.94	0.0344
BC	0.5625	1	0.5625	0.53	—
BD	0.5625	1	0.5625	0.53	—
CD	0.0625	1	0.0625	0.06	—
Error ($ABC + ABD + ACD + BCD$)	4.2500	4	1.0625		
Total	84.9375	15			

Es fácil verificar que los cuatro bloques son

Bloque 1	Bloque 2	Bloque 3	Bloque 4
$L_1 = 0, L_2 = 0$	$L_1 = 1, L_2 = 0$	$L_1 = 0, L_2 = 1$	$L_1 = 1, L_2 = 1$
(1)	*a*	*b*	*ab*
ac	*c*	*abc*	*bc*
bd	*abd*	*d*	*ad*
abcd	*bcd*	*acd*	*cd*

Este procedimiento general puede extenderse para confundir el diseño 2^k en 2^p bloques, donde $p < k$. Se empieza seleccionando los p efectos que van a confundirse, de tal modo que ninguno de los efectos elegidos sea una interacción generalizada de los demás. Entonces pueden construirse los bloques a partir de los p contrastes de definición $L_1, L_2, ..., L_p$ que están asociados con estos efectos. Además de los p efectos elegidos para confundirlos, exactamente $2^p - p - 1$ efectos adicionales se confunden con los bloques; se trata de las interacciones generalizadas de los p efectos originales elegidos. Deberá tenerse cuidado para no confundir efectos de interés potencial.

Para mayores detalles sobre confusión en el diseño factorial 2^k, consulte Montgomery (capítulo 7). Este libro contiene las pautas de selección de factores para confundirlos con bloques, de tal modo que los efectos principales y las interacciones de orden inferior no se confundan. En particular, el libro contiene una tabla de esquemas de confusión sugeridos para diseños que incluyen hasta siete factores y un rango de tamaños de los bloques, algunos de los cuales incluyen tan sólo dos corridas.

EJERCICIOS DE LA SECCIÓN 13-7

13-23. Considérense los datos de la primera réplica del ejercicio 13-13. Suponga que estas observaciones no pueden correrse todas bajo las mismas condiciones. Establezca un diseño para correr estas observaciones en dos bloques de cuatro observaciones cada uno, con *ABC* confundido. Analice los datos.

13-24. Considérense los datos de la primera réplica del ejercicio 13-14. Construya un diseño con dos bloques de ocho observaciones cada uno, con *ABCD* confundido. Analice los datos.

13-25. Repita el ejercicio 13-24 suponiendo que se necesitan cuatro bloques. Confunda *ABD* y *ABC* (y por consiguiente *CD*) con los bloques.

13-26. Construya un diseño 2^5 en dos bloques. Seleccione la interacción *ABCDE* para que se confunda con los bloques.

13-27. Construya un diseño 2^5 en cuatro bloques. Seleccione los efectos por confundir apropiados para que el mayor número posible de interacciones se confunda con los bloques.

13-28. Considérense los datos del ejercicio 13-19. Construya el diseño que se hubiera usado para llevar a cabo este experimento en dos bloques de ocho corridas cada uno. Analice los datos y saque conclusiones.

13-29. En un artículo de *Industrial and Engineering Chemistry* ("Experimentos factoriales en estudios de plantas piloto", pp. 1300-1306) se reporta un experimento para investigar el efecto de la temperatura (*A*), la producción de gas (*B*) y la concentración (*C*) sobre la concentración de la solución del producto en una unidad de recirculación.

Se usaron dos bloques con ABC confundido, y el experimento se repitió dos veces. Los datos son los siguientes:

Réplica 1

Bloque 1	Bloque 2
$(1) = 99$	$a = 18$
$ab = 52$	$b = 51$
$ac = 42$	$c = 108$
$bc = 95$	$abc = 35$

Réplica 2

Bloque 3	Bloque 4
$(1) = 46$	$a = 18$
$ab = 47$	$b = 62$
$ac = 22$	$c = 104$
$bc = 67$	$abc = 36$

a) Analice los datos de este experimento.
b) Grafique los residuales en papel probabilidad normal y contra los valores predichos. Comente las gráficas obtenidas.
c) Comente la eficiencia de este diseño. Obsérvese que el experimento se ha replicado dos veces, pero aún no se cuenta con información de la interacción ABC.
d) Sugiera un diseño mejor, específicamente uno que proporcione alguna información de *todas* las interacciones.

13-30. Considérese el diseño factorial 2^6. Establezca un diseño que se realice en cuatro bloques de 16 corridas cada uno. Demuestre que un diseño que confunde tres de las interacciones de cuatro factores con los bloques es el mejor arreglo posible de la formación de bloques.

13-8 RÉPLICA FRACCIONADA DEL DISEÑO 2^k

Cuando el número de factores de un diseño factorial 2^k se incrementa, el número de corridas necesarias aumenta con rapidez. Por ejemplo, un diseño 2^5 requiere 32 corridas. En este diseño, sólo 5 grados de libertad corresponden a los efectos principales, y 10 grados de libertad corresponden a las interacciones de dos factores. Dieciséis de los 31 grados de libertad se usan para estimar las interacciones de orden superior; es decir, las interacciones de tres factores y de órdenes superiores. Con frecuencia, estas interacciones de orden superior son de escaso interés, en particular cuando se inicia el estudio de un proceso o sistema. Si puede suponerse que ciertas interacciones de orden superior son insignificantes, entonces puede usarse un **diseño factorial fraccionado** que comprende un número de corridas menor que 2^k para obtener información sobre los efectos principales y las interacciones de orden inferior. En esta sección se introducirán las réplicas fraccionadas del diseño 2^k.

Uno de los usos principales de los diseños factoriales fraccionados es en los *experimentos tamiz* o *de selección*. Se trata de experimentos en los que se consideran muchos factores con el fin de identificar aquellos factores (en caso de haberlos) que tienen efectos grandes. Los experimentos de tamiz suelen realizarse en las etapas iniciales de un proyecto, cuando es probable que muchos de los factores considerados inicialmente tengan un efecto escaso o nulo sobre la respuesta. Después, los factores que se identifican como importantes se investigan con mayor profundidad en experimentos subsecuentes.

13-8.1 Media fracción del diseño 2^k

Una media fracción del diseño 2^k contiene 2^{k-1} corridas y suele llamarse diseño factorial fraccionado 2^{k-1}. Como un ejemplo, considérese el diseño 2^{3-1}, esto es una media fracción del diseño 2^3. Este diseño sólo tiene cuatro corridas, en contraste con el diseño factorial completo que

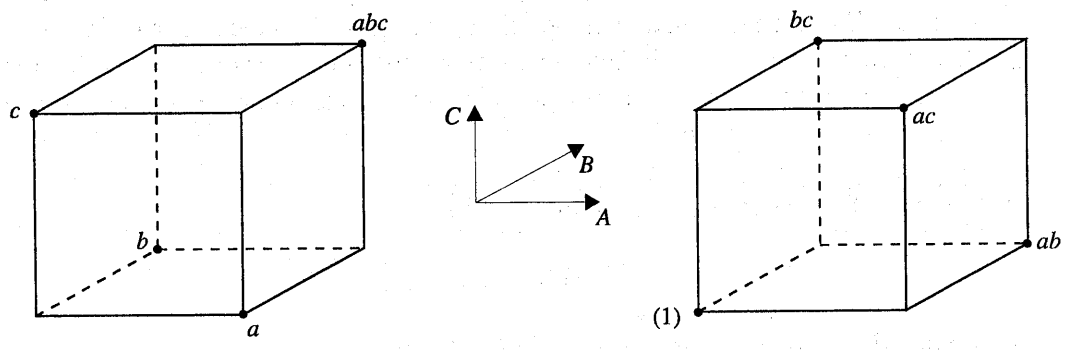
Tabla 13-26 Signos positivos y negativos del diseño factorial 2^3

Combinación de tratamientos	Efecto factorial							
	I	*A*	*B*	*C*	*AB*	*AC*	*BC*	*ABC*
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
c	+	−	−	+	+	−	−	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+	−	+	−	−	−
ac	+	+	−	+	−	+	−	−
bc	+	−	+	+	−	−	+	−
(1)	+	−	−	−	+	+	+	−

requeriría ocho corridas. La tabla de signos positivos y negativos del diseño 2^3 se muestra en la tabla 13-26.

Suponga que se seleccionan las cuatro combinaciones de tratamientos *a, b, c* y *abc* como una media fracción. Estas combinaciones de tratamientos se muestran en la mitad superior de la tabla 13-26 y en la figura 13-32*a*. Seguirá usándose tanto la notación de minúsculas (*a, b, c, ...*) como la notación geométrica o de signos positivos y negativos para las combinaciones de tratamientos. La equivalencia entre las dos notaciones es la siguiente:

Notación 1	Notación 2
a	+ − −
b	− + −
c	− − +
abc	+ + +



La fracción principal, $I = +ABC$

(*a*)

La fracción alternativa, $I = -ABC$

(*b*)

Figura 13-32 Las medias fracciones del diseño 2^3. ***a***) La fracción principal, $I = +ABC$. ***b***) La fracción alternativa, $I = -ABC$.

Obsérvese que el diseño 2^{3-1} se forma seleccionando sólo las combinaciones de tratamientos que tienen signo positivo en el efecto ABC. Por tanto, a ABC se le llama el **generador** de esta fracción particular. Además, el elemento identidad I también tiene signo positivo para las cuatro corridas, por lo que a

$$I = ABC$$

se le llama la **relación de definición** del diseño.

Las combinaciones de tratamientos de los diseños 2^{3-1} producen tres grados de libertad asociados con los efectos principales. En la mitad superior de la tabla 13-26 se obtienen las estimaciones de los efectos principales como combinaciones de las observaciones, es decir,

$$A = \tfrac{1}{2}\left[a - b - c + abc\right]$$
$$B = \tfrac{1}{2}\left[-a + b - c + abc\right]$$
$$C = \tfrac{1}{2}\left[-a - b + c + abc\right]$$

También es sencillo verificar que las estimaciones de las interacciones de dos factores deberán ser las siguientes combinaciones lineales de las observaciones:

$$BC = \tfrac{1}{2}\left[a - b - c + abc\right]$$
$$AC = \tfrac{1}{2}\left[-a + b - c + abc\right]$$
$$AB = \tfrac{1}{2}\left[-a - b + c + abc\right]$$

Por lo tanto, la combinación lineal de las observaciones de la columna A, ℓ_A, estima tanto el efecto principal de A como la interacción BC. Es decir, la combinación lineal ℓ_A estima la suma de estos dos efectos, $A + BC$. De manera similar, ℓ_B estima $B + AC$, y ℓ_C estima $C + AB$. Dos o más efectos que tienen esta propiedad reciben el nombre de **alias**. En el diseño 2^{3-1} anterior, A y BC son alias, B y AC son alias, y C y AB son alias. Los alias son resultado directo de la réplica fraccionada. En muchas situaciones prácticas, será posible seleccionar la fracción de tal modo que los efectos principales y las interacciones de orden inferior que sean de interés sólo tengan como alias las interacciones de orden superior (las cuales probablemente son insignificantes).

La estructura de los alias para este diseño se encuentra utilizando la relación de definición $I = ABC$. Al multiplicar cualquier efecto por la relación de definición, se obtiene el alias de dicho efecto. En el ejemplo tratado aquí, el alias de A es

$$A = A \cdot ABC = A^2BC = BC$$

puesto que $A \cdot I = A$ y $A^2 = I$. Los alias de B y C son

$$B = B \cdot ABC = AB^2C = AC$$

y

$$C = C \cdot ABC = ABC^2 = AB$$

Suponga ahora que se ha elegido la otra media fracción; es decir, las combinaciones de tratamientos de la tabla 13-26 asociadas con el signo negativo de ABC. Estas cuatro corridas se muestran en la mitad inferior de la tabla 13-26 y en la figura 13-32*b*. La relación de definición de este diseño es $I = -ABC$. Los alias son $A = -BC$, $B = -AC$ y $C = -AB$. Por tanto, las estimaciones de A, B y C que resultan de esta fracción en realidad estiman $A - BC$, $B - AC$ y $C - AB$. En la práctica, por lo general es indiferente la media fracción que se seleccione. A la fracción con el signo positivo en la relación de definición suele llamársele la **fracción principal**, y a la otra fracción por lo general se le llama la **fracción alternativa**.

Obsérvese que si se hubiera escogido AB como generador del diseño factorial fraccionado, entonces

$$A = A \cdot AB = B$$

y los dos efectos principales de A y B serían alias. Con esto por lo general se pierde información importante.

En ocasiones se usan **secuencias** de diseños factoriales fraccionados para estimar los efectos. Por ejemplo, supóngase que la fracción principal del diseño 2^{3-1} se corrió con el generador ABC. A partir de este diseño se obtienen las siguientes estimaciones de los efectos:

$$\ell_A = A + BC$$
$$\ell_B = B + AC$$
$$\ell_C = C + AB$$

Suponga que en este punto uno está dispuesto a considerar que las interacciones de dos factores son insignificantes. Si lo son, entonces el diseño 2^{3-1} ha producido las estimaciones de los tres efectos principales A, B y C. Sin embargo, si después de correr la fracción principal no se está seguro de las interacciones, es posible estimarlas corriendo la fracción *alternativa*. La fracción alternativa produce las siguientes estimaciones de los efectos:

$$\ell'_A = A - BC$$
$$\ell'_B = B - AC$$
$$\ell'_C = C - AB$$

Ahora es posible obtener las estimaciones, sin alias, de los efectos principales y de las interacciones de dos factores sumando y restando las combinaciones lineales de los efectos estimados en las dos fracciones individuales. Por ejemplo, suponga que se quiere separar A del alias de la interacción de dos factores BC. Puesto que $\ell_A = A + BC$ y $\ell'_A = A - BC$, las estimaciones de estos efectos pueden combinarse de la siguiente manera:

$$\frac{1}{2}\left(\ell_A + \ell'_A\right) = \frac{1}{2}\left(A + BC + A - BC\right) = A$$

y

$$\frac{1}{2}\left(\ell_A + \ell'_A\right) = \frac{1}{2}\left(A + BC - A + BC\right) = BC$$

Para los tres pares de estimaciones de los efectos se obtendrían los siguientes resultados:

Efecto, i	**de** $\frac{1}{2}(\ell_i + \ell_i')$	**de** $\frac{1}{2}(\ell_i + \ell_i')$
$i = A$	$\frac{1}{2}(A+BC+A-BC)=A$	$\frac{1}{2}[A+BC-(A-BC)]=BC$
$i = B$	$\frac{1}{2}(B+AC+B-AC)=B$	$\frac{1}{2}[B+AC-(B-AC)]=AC$
$i = C$	$\frac{1}{2}(C+AB+C-AB)=C$	$\frac{1}{2}[C+AB-(C-AB)]=AB$

Por tanto, al combinar una secuencia de dos diseños factoriales fraccionados, pueden aislarse tanto los efectos principales como las interacciones de dos factores. Esta propiedad hace que el diseño factorial fraccionado sea de gran utilidad en problemas experimentales, ya que permiten correr secuencias de experimentos pequeños y eficientes, combinar información de *varios* experimentos, y sacar provecho del conocimiento del proceso con el que se está experimentando conforme se va avanzando. Esta es una ilustración del concepto de experimentación secuencial.

Puede construirse un diseño 2^{k-1} escribiendo las combinaciones de tratamientos del diseño factorial completo con $k-1$ factores, llamado el **diseño básico**, y agregando después el factor k, identificando sus niveles positivo y negativo con los signos positivo y negativo de la interacción del orden más alto. Por lo tanto, un diseño factorial fraccionado 2^{3-1} se construye escribiendo el diseño básico como un diseño factorial 2^2 completo y después igualando el factor C con la interacción $\pm AB$. Por tanto, para construir la fracción principal, se usaría $C = +AB$ de la siguiente manera:

Diseño básico		**Diseño fraccionado**		
Completo 2^2		**$2^{3-1}, I = +ABC$**		
A	B	A	B	$C = AB$
−	−	−	−	+
+	−	+	−	−
−	+	−	+	−
+	+	+	+	+

Para obtener la fracción alternativa, se igualaría la última columna con $C = -AB$.

EJEMPLO 13-11

Para ilustrar el uso de una media fracción, considérese el experimento del grabado químico con plasma descrito en el ejemplo 13-8. Suponga que se decide usar un diseño 2^{4-1} con $I = ABCD$ para investigar los cuatro factores: entrehierro (A), presión (B), flujo de C_2F_6 (C) y ajuste de la energía (D). Este diseño se construiría escribiendo como diseño básico un diseño 2^3 con los factores A, B y C y estableciendo después los niveles del cuarto factor $D = ABC$. En la tabla 13-27 se muestra el diseño y los valores de rapidez de grabado resultantes. En la figura 13-33 se ilustra gráficamente el diseño.

Tabla 13-27 El diseño 2^{4-1} con relación de definición $I = ABCD$

A	B	C	$D = ABC$	Combinación de tratamientos	Rapidez de grab. quím.
−	−	−	−	(1)	550
+	−	−	+	*ad*	749
−	+	−	+	*bd*	1 052
+	+	−	−	*ab*	650
−	−	+	+	*cd*	1 075
+	−	+	−	*ac*	642
−	+	+	−	*bc*	601
+	+	+	+	*abcd*	729

En este diseño, los efectos principales y las interacciones de tres factores forman alias; obsérvese que el alias de A es

$$A \cdot I = A \cdot ABCD$$
$$A = A^2BCD = BCD$$

y de manera similar, $B = ACD$, $C = ABD$ y $D = ABC$.

Las interacciones de dos factores forman alias entre sí. Por ejemplo, el alias de AB es CD:

$$AB \cdot I = AB \cdot ABCD$$
$$AB = A^2B^2CD = CD$$

Los otros alias son $AC = BD$ y $AD = BC$.

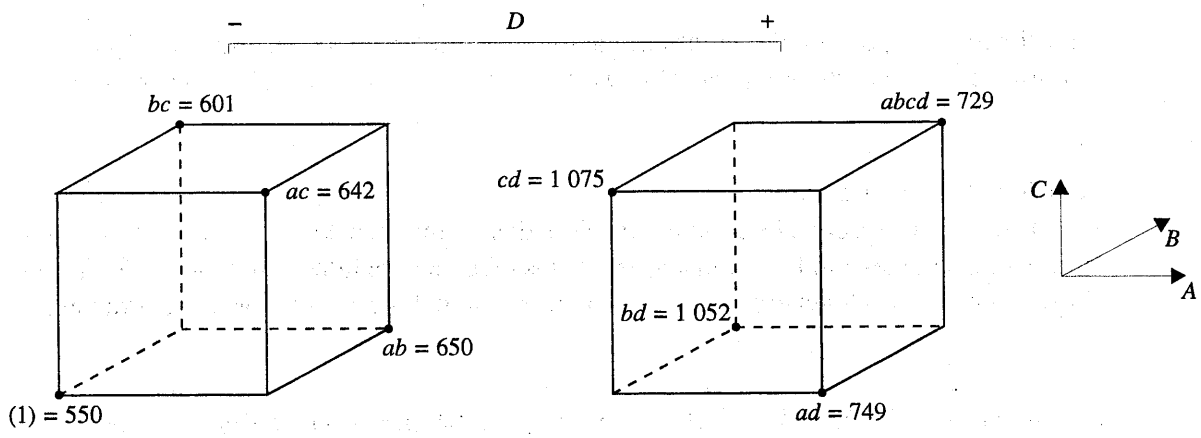


Figura 13-33 El diseño 2^{4-1} para el experimento del ejemplo 13-11.

Las estimaciones de los efectos principales y sus alias se encuentran utilizando las cuatro columnas de signos de la tabla 13-27. Por ejemplo, por la columna A se obtiene el efecto estimado

$$\ell_A = A + BCD = \tfrac{1}{4}(-550 + 749 - 1\,052 + 650 - 1\,075 + 642 - 601 + 729)$$
$$= -127.00$$

Las otras columnas producen

$$\ell_B = B + ACD = 4.00$$
$$\ell_C = C + ABD = 11.50$$

y

$$\ell_D = D + ABC = 290.50$$

Evidentemente, ℓ_A y ℓ_D son grandes, y si se piensa que las interacciones de tres factores son despreciables, entonces los efectos principales A (entrehierro) y D (ajuste de la energía) afectan significativamente la rapidez de grabado.

Las interacciones se estiman formando las columnas AB, AC y AD y agregándolas a la tabla. Por ejemplo, los signos de la columna AB son +, –, –, +, +, –, –, +, y esta columna produce la estimación

$$\ell_A = AB + CD = \tfrac{1}{4}(550 - 749 - 1\,052 + 650 + 1\,075 - 642 - 601 + 729)$$
$$= -10.00$$

A partir de las columnas AC y AD se encuentra

$$\ell_{AC} = AC + BD = -25.50$$
$$\ell_{AD} = AD + BC = -197.50$$

La estimación ℓ_{AD} es grande; la interpretación más directa de los resultados es que como A y D son grandes, se trata de la interacción AD. Por tanto, los resultados obtenidos con el diseño 2^{4-1} concuerdan con los resultados del diseño factorial completo del ejemplo 13-9.

Solución de computadora

Por lo general los diseños factoriales fraccionados se analizan con un paquete de software. En la tabla 13-28 se muestran las estimaciones de los efectos obtenidas con el programa Minitab para el ejemplo 13-11. Concuerdan con los cálculos manuales presentados anteriormente.

Gráfica de probabilidad normal de los efectos

La gráfica de probabilidad normal es muy útil para determinar la significación de los efectos de un diseño factorial fraccionado, en particular cuando son muchos los efectos por estimar. Se recomienda encarecidamente examinar esta gráfica. En la figura 13-34 se presenta la gráfica

Tabla 13-28 Estimaciones de los efectos del programa Minitab, ejemplo 13-11

Ajuste factorial fraccionado		
Efectos estimados y coeficientes de la rapidez de grabado químico		
Término	Efecto	Coeficiente
Constante		756.00
Entrehierro	−127.00	−63.50
Presión	4.00	2.00
F.	11.50	5.75
Energía	290.50	145.25
Entrehierro*Presión	−10.00	−5.00
Entrehierro*F.	−25.50	−12.75
Entrehierro*Energía	−197.50	−98.75

de probabilidad normal de los efectos del ejemplo 13-11. Esta gráfica se obtuvo con el programa Minitab. Obsérvese que *A*, *D* y la interacción *AD* destacan claramente en esta gráfica.

Análisis de residuales

Los residuales de un diseño factorial fraccionado pueden obtenerse con el método del modelo de regresión presentado anteriormente. Obsérvese que la salida de Minitab de la tabla 13-28 para el ejemplo 13-11 muestra los coeficientes de regresión. Los residuales deberán analizarse gráficamente como ya se comentó, tanto para determinar la validez de los supuestos fundamentales del modelo como para profundizar en la situación experimental.

Proyección de diseños 2^{k-1}

Si uno o más factores de una media fracción de un diseño 2^k pueden descartarse, el diseño se proyectará en un diseño factorial completo. Por ejemplo, en la figura 13-35 se presenta un

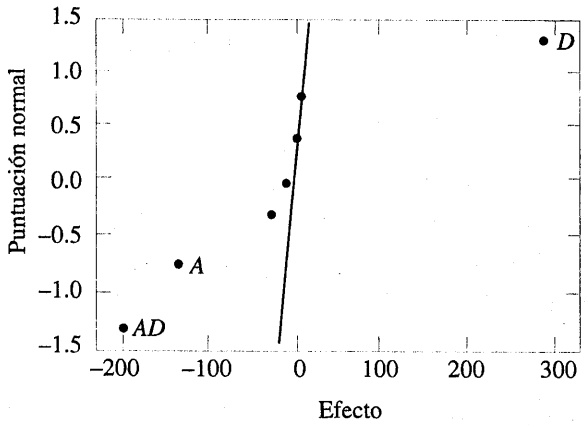


Figura 13-34 Gráfica de probabilidad normal del programa Minitab de los efectos, ejemplo 13-11.

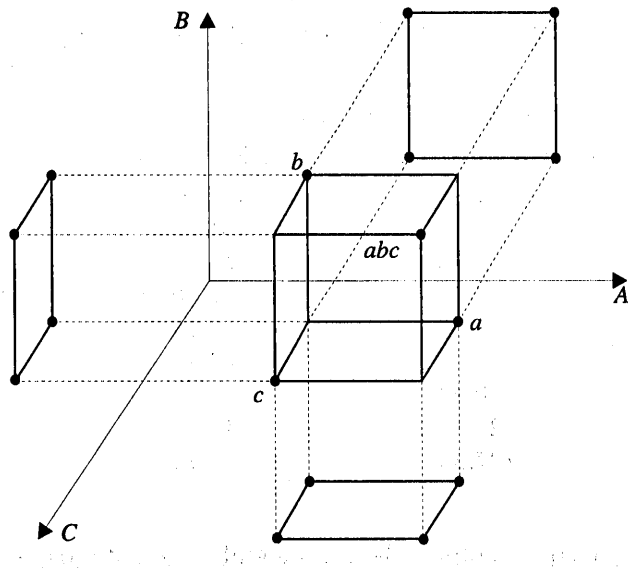


Figura 13-35 Proyección de un diseño 2^{3-1} en tres diseños 2^2.

diseño 2^{3-1}. Obsérvese que este diseño se proyectará en un diseño factorial completo con dos de los tres factores originales. Por tanto, si se piensa que a lo sumo dos de los tres factores son importantes, el diseño 2^{3-1} es excelente para identificar los factores significativos. Esta **propiedad de proyección** es muy útil en el tamizado de factores, ya que permite eliminar los factores insignificantes, dando como resultado un experimento más sólido en los factores activos que quedan.

En el diseño 2^{4-1} usado en el experimento del grabado químico con plasma del ejemplo 13-11, se encontró que dos de los cuatro factores (B y C) podrían descartarse. Si se eliminan estos dos factores, las columnas que quedan en la tabla 13-27 forman un diseño 2^2 en los factores A y D, con dos réplicas. Este diseño se muestra en la figura 13-36. Los efectos principales de A y D y la fuerte interacción de dos factores AD son claramente evidentes en esta gráfica.

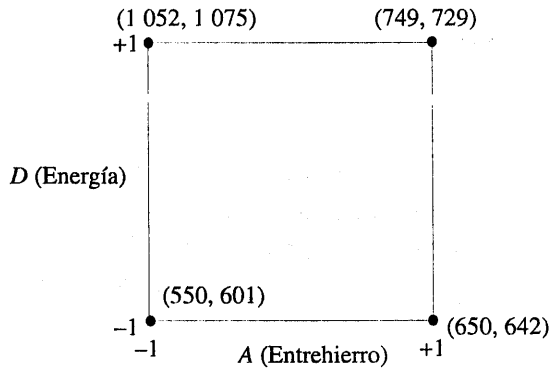


Figura 13-36 El diseño 2^2 obtenido al descartar los factores B y C del experimento del grabado químico con plasma del ejemplo 13-11.

Resolución del diseño

El concepto de resolución del diseño es un recurso útil para catalogar los diseños factoriales fraccionados de acuerdo con los patrones de los alias que producen. Los diseños de resolución III, IV y V son de particular importancia. Se presentan a continuación las definiciones de estos términos y un ejemplo de cada uno.

1. **Diseños de resolución III.** Se trata de diseños en los que ninguno de los efectos principales es alias de ningún otro efecto principal, pero los efectos principales y las interacciones de dos factores forman alias, y algunas de las interacciones de dos factores pueden ser alias entre sí. El diseño 2^{3-1} con $I = ABC$ es un diseño de resolución III. Se acostumbra usar un subíndice en números romanos para indicar la resolución del diseño; por tanto, esta media fracción es un diseño 2_{III}^{3-1}.
2. **Diseños de resolución IV.** Se trata de diseños en los que ninguno de los efectos principales es alias de ningún otro efecto principal ni de las interacciones de dos factores, pero las interacciones de dos factores forman alias entre sí. El diseño 2^{4-1} con $I = ABCD$ usado en el ejemplo 13-11 es un diseño de resolución IV $\left(2_{\text{IV}}^{4-1}\right)$.
3. **Diseños de resolución V.** Se trata de diseños en los que ninguno de los efectos principales ni de las interacciones de dos factores forman alias de otro efecto principal o interacción de dos factores, pero las interacciones de dos factores forman alias de las interacciones de tres factores. Un diseño 2^{5-1} con $I = ABCDE$ es un diseño de resolución $\left(2_{\text{V}}^{5-1}\right)$.

Los diseños de resolución III y IV son particularmente útiles en experimentos de tamizado de factores. Un diseño de resolución IV proporciona información adecuada acerca de los efectos principales y proporcionará cierta información acerca de las interacciones de dos factores.

13-8.2 Fracciones menores: el factorial fraccionado 2^{k-p}

Aun cuando el diseño 2^{k-1} es útil para reducir el número de corridas requeridas en un experimento, con frecuencia se encuentra que fracciones menores proporcionarán casi la misma información útil con una economía aún mayor. En general, un diseño 2^k puede procesarse en una fracción $1/2^p$ llamada diseño factorial fraccionado 2^{k-p}. Por tanto, a una fracción 1/4 se le llama un diseño 2^{k-2}, a una fracción 1/8 se le llama diseño 2^{k-3}, a una fracción 1/16 diseño 2^{k-4}, y así sucesivamente.

Para ilustrar la fracción 1/4, considérese un experimento con seis factores y suponga que el interés principal del ingeniero está en los efectos principales, pero le gustaría tener también alguna información sobre las interacciones de dos factores. Un diseño 2^{6-1} requeriría 32 corridas y tendría 31 grados de libertad para estimar los efectos. Puesto que sólo hay 6 efectos principales y 15 interacciones de dos factores, la media fracción no es eficiente, requiere demasiadas corridas. Suponga que se considera una fracción 1/4, es decir, un diseño 2^{6-2}. Este diseño contiene 16 corridas y, con 15 grados de libertad, permitirá que se estimen los 6 efectos principales, con cierto margen para examinar las interacciones de dos factores.

Para generar este diseño, se escribiría un diseño 2^4 con los factores A, B, C y D como diseño básico, y después se agregarían las columnas E y F. Para encontrar las nuevas columnas podrían seleccionarse los dos **generadores del diseño** $I = ABCE$ e $I = BCDF$. Por tanto, la columna E se encontraría a partir de $E = ABC$, y la columna F sería $F = BCD$. Es decir, las columnas $ABCE$ y $BCDF$ son iguales a la columna identidad. Sin embargo, se sabe que el productos de dos colum-

nas cualesquiera de la tabla de signos positivos y negativos para un diseño 2^k es tan sólo otra columna de la tabla; por lo tanto, el producto de $ABCE$ y $BCDF$ o $ABCE(BCDF) = AB^2C^2DEF = ADEF$ es también una columna identidad. Por consiguiente, la **relación de definición completa** del diseño 2^{6-2} es

$$I = ABCE = BCDF = ADEF$$

Se hace referencia a cada término de una relación de definición (tal como $ABCE$ arriba) como una **palabra**. Para encontrar el alias de cualquiera de los efectos, simplemente se multiplica el efecto por cada palabra de la relación de definición anterior. Por ejemplo, el alias de A es

$$A = BCE = ABCDF = DEF$$

En la tabla 13-29 se muestran las relaciones completas de los alias para este diseño. En general, la resolución de un diseño 2^{k-p} es igual al número de letras incluidas en la palabra más corta de la relación de definición completa. Por lo tanto, se trata de un diseño de resolución IV; los efectos principales son alias de las interacciones de tres factores y de las interacciones de órdenes superiores, y las interacciones de dos factores son alias entre sí. Este diseño proporcionaría información adecuada sobre los efectos principales y daría cierta idea acerca de la fuerza de las interacciones de dos factores. La construcción y el análisis del diseño se ilustran en el ejemplo 13-12.

EJEMPLO 13-12

La piezas manufacturadas en un proceso de moldeo por inyección están presentando una contracción excesiva, lo cual está ocasionando problemas en las operaciones de ensamblaje en las áreas posteriores al moldeo por inyección. En un intento por reducir la contracción, un equipo de mejoramiento de la calidad ha decidido usar un experimento diseñado para estudiar el proceso de moldeo por inyección. El equipo investiga seis factores —temperatura de moldeo (A), velocidad del enroscado (B), tiempo de retención (C), duración del ciclo (D), tamaño del vaciadero (E), y la presión de retención (F)— con dos niveles cada uno, con el fin de saber cómo afecta cada factor la contracción y de obtener información preliminar acerca de la forma en que interactúan los factores.

El equipo decide usar un diseño factorial fraccionado de 16 corridas con dos niveles para estos seis factores. El diseño se construye escribiendo un diseño 2^4 como diseño básico con los

Tabla 13-29 Estructura de los alias del diseño 2_{IV}^{6-2} con $I = ABCE = BCDF = ADEF$

$A = BCE = DEF = ABCDF$	$AB = CE = ACDF = BDEF$
$B = ACE = CDF = ABDEF$	$AC = BE = ABDF = CDEF$
$C = ABE = BDF = ACDEF$	$AD = EF = BCDE = ABCF$
$D = BCF = AEF = ABCDE$	$AE = BC = DF = ABCDEF$
$E = ABC = ADF = BCDEF$	$AF = DE = BCEF = ABCD$
$F = BCD = ADE = ABCEF$	$BD = CF = ACDE = ABEF$
$ABD = CDE = ACF = BEF$	$BF = CD = ACEF = ABDE$
$ACD = BDE = ABF = CEF$	

Tabla 13-30 Diseño 2_{IV}^{6-2} para el experimento del moldeo por inyección del ejemplo 13-12

Corrida	A	B	C	D	$E = ABC$	$F = BCD$	Contracción observada (× 10)
1	−	−	−	−	−	−	6
2	+	−	−	−	+	−	10
3	−	+	−	−	+	+	32
4	+	+	−	−	−	+	60
5	−	−	+	−	+	+	4
6	+	−	+	−	−	+	15
7	−	+	+	−	−	−	26
8	+	+	+	−	+	−	60
9	−	−	−	+	−	+	8
10	+	−	−	+	+	+	12
11	−	+	−	+	+	−	34
12	+	+	−	+	−	−	60
13	−	−	+	+	+	−	16
14	+	−	+	+	−	−	5
15	−	+	+	+	−	+	37
16	+	+	+	+	+	+	52

factores A, B, C y D y haciendo $E = ABC$ y $F = BCD$, como se señaló antes. En la tabla 13-30 se muestra el diseño, junto con la contracción observada (× 10) de la pieza de prueba producida en cada una de las 16 corridas del diseño.

En la figura 13-37 se muestra una gráfica de probabilidad normal de las estimaciones de los efectos de este experimento. Los únicos efectos grandes son A (temperatura de moldeo), B (velocidad del enroscado) y la interacción AB. Considerando la relación de los alias de la tabla 13-29, parece razonable adoptar estas conclusiones de manera tentativa. La gráfica de la interacción AB de la figura 13-38 indica que el proceso no es sensible a la temperatura si la velocidad del enroscado está en el nivel bajo, pero que es sensible a la temperatura si la velocidad del enroscado está en el nivel alto. Con la velocidad del enroscado en un nivel bajo, el proceso deberá producir una contracción promedio de cerca de 10% independientemente del nivel de temperatura escogido.

Con base en este análisis inicial, el equipo decide fijar la temperatura de moldeo y la velocidad del enroscado en el nivel bajo. Este ajuste de las condiciones debería reducir la contracción media de las piezas aproximadamente 10%. Sin embargo, la variabilidad de la contracción de una pieza a otra sigue siendo un problema potencial. De hecho, la contracción media puede reducirse adecuadamente mediante las modificaciones anteriores; sin embargo, la variabilidad de una pieza a otra en una corrida de producción podría seguir causando problemas en el ensamblaje. Una forma de abordar esta cuestión es ver si alguno de los factores del proceso afecta la variabilidad de la contracción de las piezas.

En la figura 13-39 se presenta la gráfica de probabilidad normal de los residuales. Esta gráfica parece ser satisfactoria. Se construyeron después las gráficas de los residuales contra cada factor. Una de estas gráficas, la de los residuales contra el factor C (tiempo de retención),

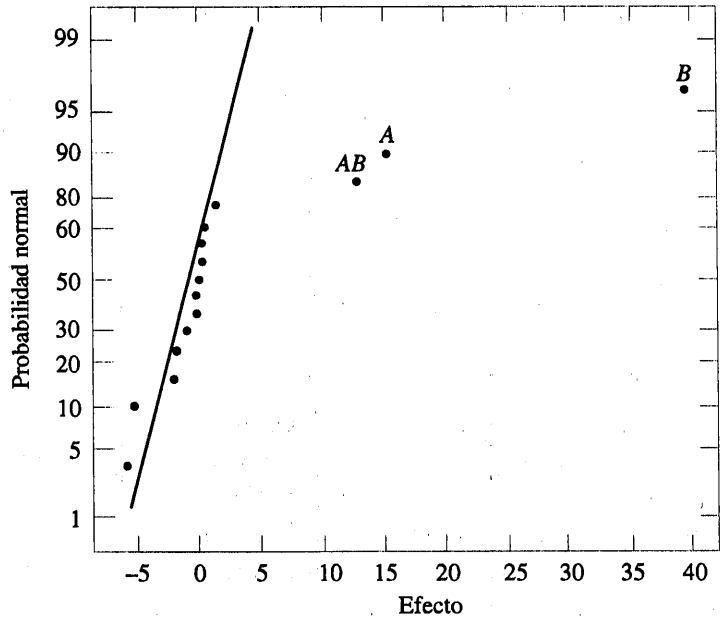


Figura 13-37 Gráfica de probabilidad normal de los efectos del ejemplo 13-12.

se muestra en la figura 13-40. La gráfica revela una dispersión mucho menor en los residuales cuando el tiempo de retención es bajo que cuando es alto. Estos residuales se obtuvieron de la manera usual a partir del modelo para la contracción predicha

$$\begin{aligned}\hat{y} &= \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_{12} x_1 x_2 \\ &= 27.3125 + 6.9375 x_1 + 17.8125 x_2 + 5.9375 x_1 x_2\end{aligned}$$

donde x_1, x_2 y x_1x_2 son variables codificadas que corresponden a los factores A y B y la interacción AB. Entonces los residuos son

$$e = y - \hat{y}$$

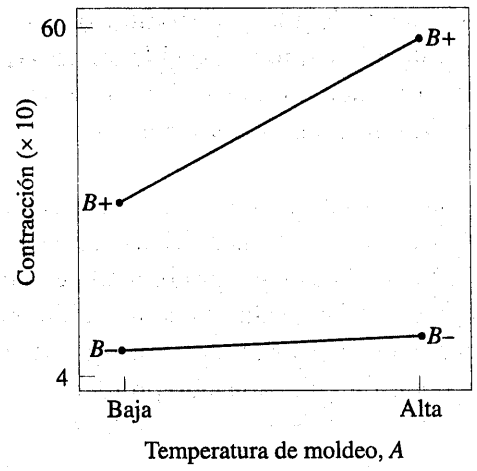


Figura 13-38 Gráfica de la interacción AB (temperatura de moldeo-velocidad del enroscado) del ejemplo 13-12.

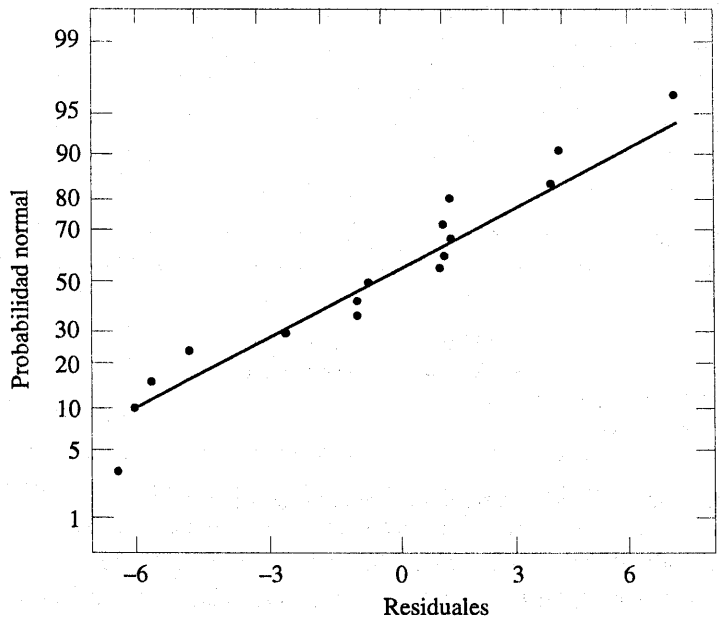


Figura 13-39 Gráfica de probabilidad normal de los residuales del ejemplo 13-12.

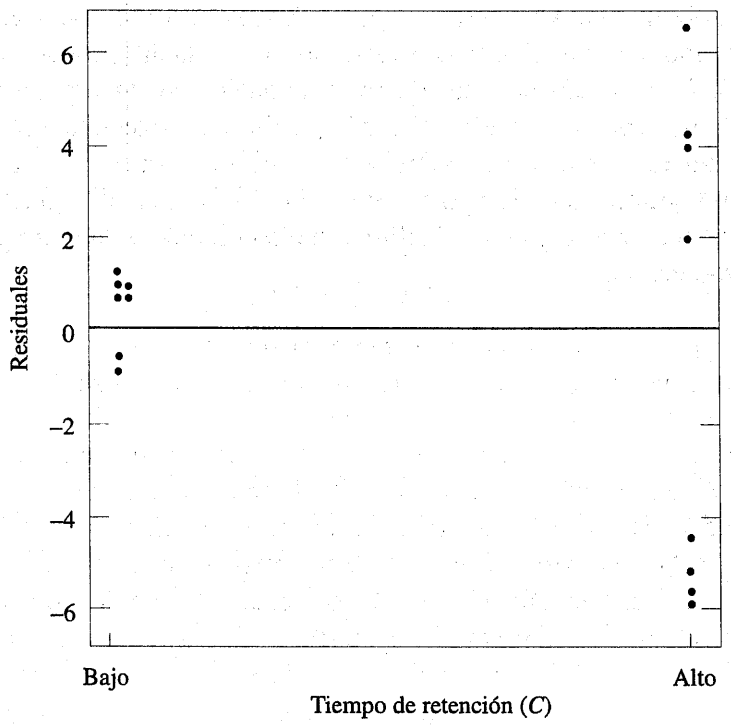


Figura 13-40 Residuales contra el tiempo de retención (C) del ejemplo 13-12.

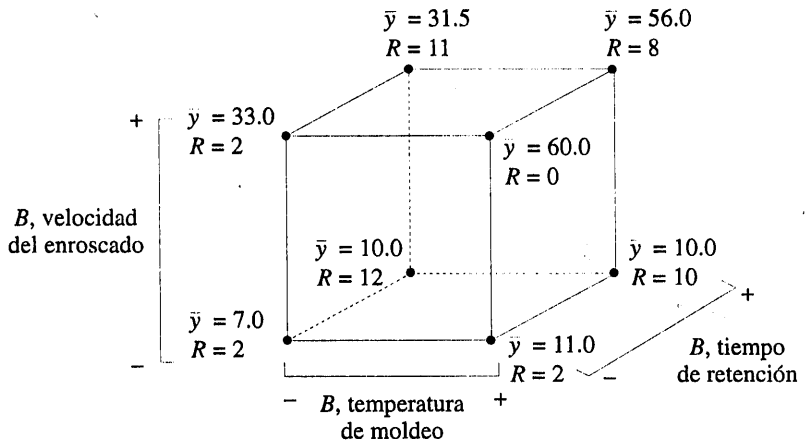


Figura 13-41 Contracción promedio y rango de la contracción con los factores A, B y C para el ejemplo 13-12.

El modelo de regresión usado para producir los residuales, básicamente elimina los efectos de localización de A, B y AB de los datos; en consecuencia, los residuales contienen información acerca de la variabilidad no explicada. La figura 13-40 indica que existe un patrón en la variabilidad y que la variabilidad de la contracción de las piezas puede ser menor cuando el tiempo de retención está en el nivel bajo.

En la figura 13-41 se muestran los datos de este experimento proyectados en un cubo con los factores A, B y C. En cada vértice del cubo se muestra la contracción promedio observada y el rango de la contracción observada. Al examinar esta figura, se observa que correr el proceso con la velocidad del enroscado (B) en el nivel bajo es la clave para reducir la contracción promedio de las piezas. Si B es baja, prácticamente cualquier combinación de la temperatura (A) y de el tiempo de retención (C) dará como resultado valores bajos de la contracción promedio de las piezas. Sin embargo, al examinar los rangos de los valores de la contracción en cada vértice del cubo, se observa de inmediato que ajustar e tiempo de retención (C) en el nivel bajo es la elección más apropiada si se quiere mantener baja la variabilidad de la contracción de una pieza a otra durante una corrida de producción.

Los conceptos usados para construir el diseño factorial fraccionado 2^{6-2} el ejemplo 13-2 pueden extenderse a la construcción de cualquier diseño factorial fraccionado 2^{k-p}. Un diseño factorial fraccionado 2^k que contiene 2^{k-p} corridas, en general se llama fracción $1/2^p$ del diseño 2^k o simplemente, diseño factorial fraccionado 2^{k-p}. Estos diseños requieren la selección de p generadores independientes. La relación de definición del diseño se compone de los p generadores elegidos inicialmente y sus $2^p - p - 1$ interacciones generalizadas.

La estructura de los alias puede encontrarse multiplicando la columna de cada efecto por la relación de definición. Deberá tenerse cuidado al elegir los generadores para que los efectos de interés potencial no formen alias entre sí. Cada efecto tiene $2^p - 1$ alias. Para valores moderadamente grandes de k, por lo general se supone que las interacciones de orden superior (por ejemplo, de tercer o cuarto orden o mayor) son insignificantes, con lo cual se simplifica en gran medida la estructura de los alias.

Es importante seleccionar los p generadores del diseño factorial fraccionado 2^{k-p} de tal modo que se obtengan las mejores relaciones de los alias posibles. Un criterio razonable es

seleccionar los generadores para que el diseño 2^{k-p} resultante tenga la resolución de diseño más alta posible. Montgomery presenta una tabla de generadores recomendados para diseños factoriales fraccionados 2^{k-p} para $k \leq 15$ factores y hasta $n \leq 128$ corridas. Una parte de esta tabla se reproduce aquí en la tabla 13-31. En esta tabla, los generadores se muestran con las opciones + o –; la selección de todos los generadores como + dará una fracción principal, mientras que si algunos de los generadores son elecciones –, el diseño será una de las fracciones alternativas de la misma familia. Los generadores sugeridos en esta tabla darán como resultado un diseño con la resolución más alta posible. Montgomery también da una tabla de las relaciones de los alias para estos diseños.

EJEMPLO 13-13

Para ilustrar el uso de la tabla 13-31, suponga que se tienen siete factores y que hay interés en estimar los siete efectos principales y en formarse una idea de las interacciones de dos factores. Uno está dispuesto a suponer que las interacciones de tres factores y de órdenes superiores son insignificantes. Esta información sugiere que un diseño de resolución IV sería apropiado.

En la tabla 13-31 se observa que hay dos fracciones de resolución IV disponibles: la 2_{IV}^{7-2} con 32 corridas y la 2_{IV}^{7-3} con 16 corridas. En la tabla 13-32 se presentan los alias que incluyen los efectos principales y las interacciones de dos y tres factores para el diseño con 16 corridas. Obsérvese que los siete efectos principales y las interacciones de tres factores forman alias. Todas las interacciones de dos factores forman alias en grupos de tres. Por lo tanto, este diseño satisfará los objetivos propuestos; es decir, permitirá la estimación de los efectos principales, y dará cierta idea de las interacciones de dos factores. No es necesario correr el diseño 2_{IV}^{7-2}, el cual requeriría 32 corridas. La construcción del diseño 2_{IV}^{7-3} se muestra en la tabla 13-33. Obsérvese que se construyó empezando con el diseño 2^4 con 16 corridas en A, B, C y D como diseño básico y agregando después las tres columnas $E = ABC$, $F = BCD$, y $G = ACD$ como se sugiere en la tabla 13-31. Por tanto, los generadores de este diseño son $I = ABCE$, $I = BCDF$ e $I = ACDG$. La relación de definición completa es $I = ABCE = BCDF = ADEF = ACDG = BDEG = CEFG = ABFG$. Esta relación de definición se usó para producir los alias de la tabla 13-32. Por ejemplo, la relación de los alias de A es

$$A = BCE = ABCDF = DEF = CDG = ABDEG = ACEFG = BFG$$

la cual, si se ignoran las interacciones con más de tres factores, concuerda con la tabla 13-32.

Para siete factores, el número de corridas puede reducirse aún más. El diseño 2^{7-4} es un experimento con ocho corridas en las que se incluyen siete variables. Se trata de una fracción 1/16 y se obtiene escribiendo primero un diseño 2^3 como diseño básico con los factores A, B, C, y formando después las cuatro nuevas columnas a partir de $I = ABD$, $I = ACE$, $I = BCF$ e $I = ABCG$, como se sugiere en la tabla 13-31. El diseño se muestra en la tabla 13-34.

La relación de definición completa se encuentra multiplicando los generadores de dos en dos, de tres en tres y, por último, de cuatro en cuatro, siendo el resultado

$$\begin{aligned} I &= ABD = ACE = BCF = ABCG = BCDE = ACDF = CDG = ABEF \\ &= BEG = AFG = DEF = ADEG = CEFG = BDFG = ABCDEFG \end{aligned}$$

Tabla 13-31 Diseños factoriales fraccionados 2^{k-p} selecionados

Número de factores k	Fracción	Número de corridas	Generadores del diseño
3	2_{III}^{3-1}	4	$C = \pm AB$
4	2_{III}^{4-1}	8	$D = \pm ABC$
5	2_{III}^{5-1}	16	$E = \pm ABCD$
	2_{III}^{5-2}	8	$D = \pm AB$
			$E = \pm AC$
6	2_{VI}^{6-1}	32	$F = \pm ABCDE$
	2_{IV}^{6-2}	16	$E = \pm ABC$
			$F = \pm BCD$
	2_{III}^{6-3}	8	$D = \pm AB$
			$E = \pm AC$
			$F = \pm BC$
7	2_{VII}^{7-1}	64	$G = \pm ABCDEF$
	2_{IV}^{7-2}	32	$E = \pm ABC$
			$G = \pm ABDE$
	2_{IV}^{7-3}	16	$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
	2_{III}^{7-4}	8	$D = \pm AB$
			$E = \pm AC$
			$F = \pm BC$
			$G = \pm ABC$
8	2_{V}^{8-2}	64	$G = \pm ABCD$
			$H = \pm ABEF$
	2_{IV}^{8-3}	32	$F = \pm ABC$
			$G = \pm ABD$
			$H = \pm BCDE$
	2_{IV}^{8-4}	16	$E = \pm BCD$
			$F = \pm ACD$
			$G = \pm ABC$
			$H = \pm ABD$
9	2_{VI}^{9-2}	128	$H = \pm ACDFG$
			$J = \pm BCEFG$
	2_{IV}^{9-3}	64	$G = \pm ABCD$
			$H = \pm ACEF$
	2_{IV}^{9-4}	32	$J = \pm CDEF$
			$F = \pm BCDE$
			$G = \pm ACDE$
			$H = \pm ABDE$
			$J = \pm ABCE$
	2_{III}^{9-5}	16	$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
			$H = \pm ABD$
			$J = \pm ABCD$

Continúa

Tabla 13-31 Diseños factoriales fraccionados 2^{k-p} selecionados (*continuación*)

Número de factores k	Fracción	Número de corridas	Generadores del diseño
10			$H = \pm ABCG$
			$J = \pm ACDE$
	2_{V}^{10-3}	128	$K = \pm ACDF$
			$G = \pm BCDF$
			$H = \pm ACDF$
			$J = \pm ABDE$
	2_{IV}^{10-4}	64	$K = \pm ABCE$
			$F = \pm ABCD$
			$G = \pm ABCE$
			$H = \pm ABDE$
			$J = \pm ACDE$
	2_{IV}^{10-5}	32	$K = \pm BCDE$
			$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
			$H = \pm ABD$
			$J = \pm ABCD$
	2_{III}^{10-6}	16	$K = \pm AB$
11			$G = \pm CDE$
			$H = \pm ABCD$
			$J = \pm ABF$
			$K = \pm BDEF$
	2_{IV}^{11-5}	64	$L = \pm ADEF$
			$F = \pm ABC$
			$G = \pm BCD$
			$H = \pm CDE$
			$J = \pm ACD$
			$K = \pm ADE$
	2_{IV}^{11-6}	32	$L = \pm BDE$
			$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
			$H = \pm ABD$
			$J = \pm ABCD$
			$K = \pm AB$
	2_{III}^{11-7}	16	$L = \pm AC$

Fuente: Montgomery.

Tabla 13-32 Generadores, relación de definición y alias del diseño factorial fraccionado 2_{IV}^{7-3}

Generadores y relación de definición

$$E = ABC, \quad F = BCD, \quad G = ACD$$

$$I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG$$

Alias

$A = BCE = DEF = CDG = BFG$	$AB = CE = FG$
$B = ACE = CDF = DEG = AFG$	$AC = BE = DG$
$C = ABE = BDF = ADG = EFG$	$AD = EF = CG$
$D = BCF = AEF = ACG = BEG$	$AE = BC = DF$
$E = ABC = ADF = BDG = CFG$	$AF = DE = BG$
$F = BCD = ADE = ABG = CEG$	$AG = CD = BF$
$G = ACD = BDE = ABF = CEF$	$BD = CF = EG$

$$ABD = CDE = ACF = BEF = BCG = AEG = DFG$$

Tabla 13-33 Diseño factorial fraccionado 2_{IV}^{7-3}

	Diseño básico						
Corrida	A	B	C	D	$E = ABC$	$F = BCD$	$G = ACD$
1	−	−	−	−	−	−	−
2	+	−	−	−	+	−	+
3	−	+	−	−	+	+	−
4	+	+	−	−	−	+	+
5	−	−	+	−	+	+	+
6	+	−	+	−	−	+	−
7	−	+	+	−	−	−	+
8	+	+	+	−	+	−	−
9	−	−	−	+	−	+	−
10	+	−	−	+	+	+	−
11	−	+	−	+	+	−	+
12	+	+	−	+	−	−	−
13	−	−	+	+	+	−	−
14	+	−	+	+	−	−	+
15	−	+	+	+	−	+	−
16	+	+	+	+	+	+	+

Tabla 13-34 Diseño factorial fraccionado 2_{III}^{7-4}

A	B	C	$D = AB$	$E = AC$	$F = BC$	$G = ABC$
−	−	−	+	+	+	−
+	−	−	−	−	+	+
−	+	−	−	+	−	+
+	+	−	+	−	−	−
−	−	+	+	−	−	+
+	−	+	−	+	−	−
−	+	+	−	−	+	−
+	+	+	+	+	+	+

El alias de cualquier efecto principal se encuentra multiplicando dicho efecto por cada término de la relación de definición. Por ejemplo, el alias de A es

$$\begin{aligned} A &= BD = CE = ABCF = BCG = ABCDE = CDF = ACDG \\ &= BEF = ABEG = FG = ADEF = DEG = ACEFG = ABDFG \\ &= BCDEFG \end{aligned}$$

Este diseño es de resolución III, ya que el efecto principal y las interacciones de dos factores forman alias. Si se supone que todas las interacciones de tres o más factores son insignificantes, los alias de los siete efectos principales son

$$\begin{aligned} \ell_A &= A + BD + CE + FG \\ \ell_B &= B + AD + CF + EG \\ \ell_C &= C + AE + BF + DG \\ \ell_D &= D + AB + CG + EF \\ \ell_E &= E + AC + BG + DF \\ \ell_F &= F + BC + AG + DE \\ \ell_G &= G + CD + BE + AF \end{aligned}$$

A este diseño 2_{III}^{7-4} se le llama **diseño factorial fraccionado saturado**, porque todos los grados de libertad disponibles se usan para estimar los efectos principales. Es posible combinar secuencias de estos diseños factoriales fraccionados de resolución III para separar los efectos principales de las interacciones de dos factores. El procedimiento se ilustra en Montgomery y en Box, Hunter y Hunter.

EJERCICIOS DE LA SECCIÓN 13-8

13-31. R.D. Snee ("Experimentación con un gran número de variables", en *Experiments in Industry: Design, Analysis and Interpretation of Results*, de R.D. Snee, L.D. Hare y J.B. Trout, eds., ASQC) describe un experimento en el que se utilizó un diseño 2^{5-1} con $I = ABCDE$ para investigar los efectos de cinco factores sobre el color de un pro-

ducto químico. Los factores son A = solvente/reactivo, B = catalizador/reactivo, C = temperatura, D = pureza del reactivo y E = pH del reactivo. Los resultados obtenidos son los siguientes:

e	=	−0.63	d	=	6.79
a	=	2.51	ade	=	6.47
b	=	−2.68	bde	=	3.45
abe	=	1.66	abd	=	5.68
c	=	2.06	cde	=	5.22
ace	=	1.22	acd	=	4.38
bce	=	−2.09	bcd	=	4.30
abc	=	1.93	$abcde$	=	4.05

a) Construya una gráfica de probabilidad normal de los efectos. ¿Cuáles son los factores activos?

b) Calcule los residuales. Construya una gráfica de probabilidad normal de los residuales y grafique éstos contra los valores ajustados. Comente las gráficas.

c) Si algunos de los factores son insignificantes, reducir el diseño 2^{5-1} a un diseño factorial completo con los factores activos. Comente el diseño resultante e interprete los resultados.

13-32. Montgomery describe un diseño factorial fraccionado 2^{4-1} usado para estudiar cuatro factores de un proceso químico. Los factores son A = temperatura, B = presión, C = concentración y D = rapidez de agitación, y la respuesta es la rapidez de filtración. El diseño y los datos son los siguientes:

Corrida	A	B	C	$D = ABC$	Combinación de tratamientos	Rapidez de filtración
1	−	−	−	−	(1)	45
2	+	−	−	+	ad	100
3	−	+	−	+	bd	45
4	+	+	−	−	ab	65
5	−	−	+	+	cd	75
6	+	−	+	−	ac	60
7	−	+	+	−	bc	80
8	+	+	+	+	$abcd$	96

a) Escriba las relaciones de los alias.

b) Estime los efectos de los factores. ¿Cuáles de los efectos de los factores parecen ser grandes?

c) Proyecte este diseño en un diseño factorial completo con los tres factores aparentemente importantes y ofrezca una interpretación práctica de los resultados.

13-33. En un artículo de *Industrial and Engineering Chemistry* ("Comentarios adicionales sobre la planeación de experimentos para aumentar la eficiencia de la investigación", pp. 60-65) se utiliza un diseño 2^{5-2} para investigar el efecto de A = temperatura de condensación, B = cantidad del material 1, C = volumen de solvente, D = tiempo de condensación y E = cantidad del material 2 sobre el rendimiento del proceso. Los resultados obtenidos son los siguientes:

e	=	23.2	cd	=	23.8
ab	=	15.5	ace	=	23.4
ad	=	16.9	bde	=	16.8
bc	=	16.2	$abcde$	=	18.1

a) Verifique que los generadores del diseño usados fueron $I = ACE$ e $I = BDE$.

b) Escriba la relación de definición completa y los alias del diseño.

c) Estime los efectos principales.

d) Construya una tabla del análisis de varianza. Verifique que las interacciones AB y AD pueden usarse como error.

e) Grafique los residuales contra los valores ajustados. Construya también una

gráfica de probabilidad normal de los residuales. Comente los resultados.

13-34. Considere el diseño 2^{6-2} de la tabla 13-30. Suponga que después de analizar los datos originales se encuentra que los factores C y E pueden descartarse. ¿Qué tipo de diseño 2^k queda para las variables restantes?

13-35. Considere el diseño 2^{6-2} de la tabla 13-30. Suponga que después de analizar los datos originales se encuentra que los factores D y F pueden descartarse. ¿Qué tipo de diseño 2^k queda para las variables restantes? Compare los resultados con el ejercicio 13-34. ¿Puede usted explicar por qué son diferentes las respuestas?

13-36. Suponga que en el ejercicio 13-14 sólo fue posible realizar una media fracción del diseño 2^4. Construya el diseño y use sólo los datos de las ocho corridas que ha generado usted para hacer el análisis.

13-37. Suponga que en el ejercicio 13-17 sólo pudo correrse una fracción un cuarto del diseño 2^5. Construya el diseño y analice los datos que se obtienen al seleccionar únicamente la respuesta de las ocho corridas de este último diseño.

13-38. Construya el diseño 2_{IV}^{8-4} recomendado en la tabla 13-31. ¿Cuáles son los alias de los efectos principales y las interacciones de dos factores?

13-39. Construya un diseño factorial fraccionado 2_{III}^{6-3}. Escriba los alias, suponiendo que sólo son de interés los efectos principales y las interacciones de dos factores.

13-40. Considérese el problema del ejercicio 13-20. Suponga que sólo pudo procesarse la mitad de las 32 corridas.

a) Elija la mitad que en su opinión debería correrse.
b) Escriba las relaciones de los alias del diseño.
c) Estime los efectos de los factores.
d) Grafique las estimaciones de los efectos en papel probabilidad normal e interprete los resultados.
e) Haga un análisis de varianza para los factores que en la gráfica de probabilidad normal del inciso d) se identificaron de interés potencial.
f) Analice los residuales del modelo.
g) Ofrezca una interpretación práctica de los resultados.

13-9 MÉTODOS Y DISEÑOS DE SUPERFICIES DE RESPUESTA

La metodología de superficies de respuesta, o MSR, es un conjunto de técnicas matemáticas y estadísticas que son herramientas de modelado y análisis útiles en aplicaciones donde una respuesta de interés recibe la influencia de varias variables y el objetivo es **optimizar esta respuesta**. Por ejemplo, suponer que un ingeniero químico quiere encontrar los niveles de la temperatura (x_1) y la concentración de alimentación (x_2) que maximicen el rendimiento (y) de un proceso. El rendimiento del proceso es una función de los niveles de la temperatura y la concentración de alimentación, por ejemplo,

$$Y = f(x_1, x_2) + \varepsilon$$

donde ε representa el ruido o error observado en la respuesta Y. Si la respuesta esperada se denota por $E(Y) = f(x_1, x_2) = \eta$, entonces a la superficie representada por

$$\eta = f(x_1, x_2)$$

se le llama **superficie de respuesta**.

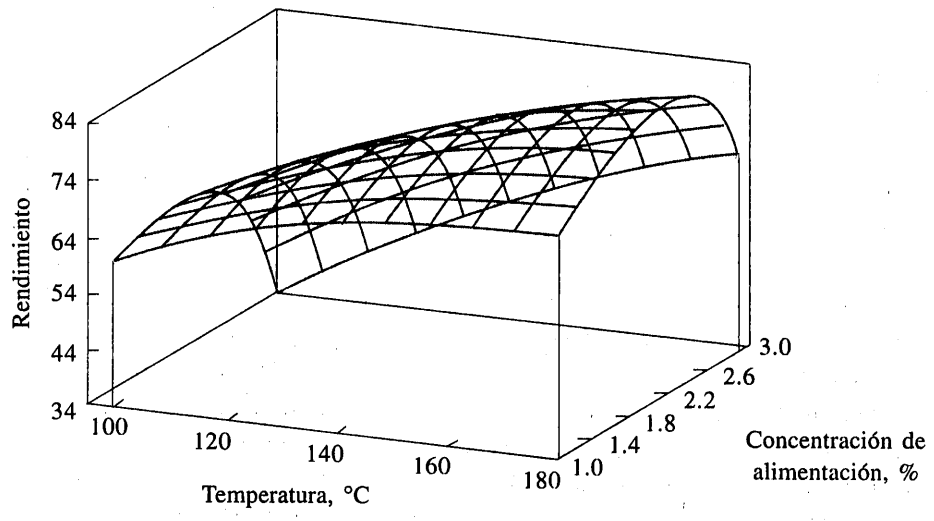


Figura 13-42 Superficie de respuesta tridimensional que indica el rendimiento esperado como una función de la temperatura y la concentración de alimentación.

La superficie de respuesta puede representarse gráficamente como se ilustra en la figura 13-42, donde η se grafica contra los niveles de x_1 y x_2. Obsérvese que la respuesta se representa como una gráfica de superficie en un espacio tridimensional. Para ayudar a visualizar la forma de una superficie de respuesta, con frecuencia se grafican los contornos de la superficie de respuesta como en la figura 13-42. En la gráfica de contorno, las líneas de respuesta constante se trazan en el plano x_1, x_2. Cada contorno corresponde a una altura particular de la superficie de respuesta. La gráfica de contorno es útil para estudiar los niveles de x_1 y x_2 que producen cambios en la forma o altura de la superficie de respuesta.

En la mayoría de los problemas MSR, se desconoce la forma de la relación entre la respuesta y las variables independientes. Por tanto, el primer paso de la MSR es encontrar una aproximación adecuada para la verdadera relación entre Y y las variables independientes. En general, se emplea un polinomio de orden inferior en alguna región de las variables independientes. Si la respuesta está bien modelada por una función lineal de las variables independientes, entonces la función de aproximación es el **modelo de primer orden**

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{13-35}$$

Si hay curvatura en el sistema, entonces debe usarse un polinomio de grado superior, como el **modelo de segundo orden**

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{\substack{i \\ i<j}} \sum_{j} \beta_{ij} x_i x_j + \varepsilon \tag{13-36}$$

En muchos problemas MSR se utilizan uno de estos polinomios de aproximación, o los dos. Por supuesto, es improbable que un modelo polinomial sea una aproximación razonable de la verda-

dera relación funcional en el espacio completo de las variables independientes, pero en general resultará adecuado para una región relativamente pequeña.

Se usa el método de mínimos cuadrados, analizado en los capítulos 10 y 11, para estimar los parámetros de los polinomios de aproximación. Después se hace el análisis de la superficie de respuesta en términos de la superficie ajustada. Si la superficie ajustada es una aproximación adecuada de la verdadera función de respuesta, entonces el análisis de la superficie ajustada será aproximadamente equivalente al análisis del sistema real.

La MSR es un procedimiento **secuencial**. Con frecuencia, cuando uno se encuentra en un punto de la superficie de respuesta que está muy apartado del óptimo (como las condiciones de operación actuales de la figura 13-43) hay poca curvatura en el sistema y el modelo de primer orden será apropiado. El objetivo en este caso es llevar al experimentador de manera rápida y eficiente a la vecindad general del óptimo. En la figura 13-43 se observa que el análisis de una superficie de respuesta puede considerarse como "el ascenso de una colina", donde la cima de la colina representa el punto de la respuesta máxima. Si el verdadero óptimo es un punto de respuesta mínima, entonces puede pensarse en "el descenso a un valle".

El objetivo final de la MSR es determinar las condiciones de operación óptimas del sistema o determinar una región del espacio de los factores en la que se satisfagan las especificaciones de operación. Asimismo, obsérvese que en la MSR la palabra "óptimo" se emplea en un sentido especial. Los procedimientos para "el ascenso de la colina" de la MSR sólo garantizan la convergencia a un óptimo local.

13-9.1 Método del ascenso más pronunciado

Frecuentemente, la estimación inicial de las condiciones de operación óptimas estarán lejos del óptimo real. En tales circunstancias, el objetivo del experimentador es pasar con rapidez a la

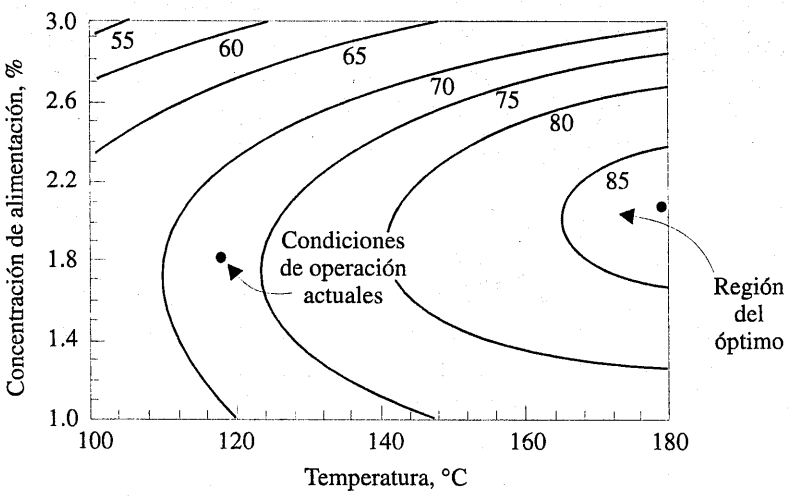


Figura 13-43 Gráfica de contorno de la superficie de respuesta del rendimiento en la figura 13-42.

vecindad general del óptimo. Se desea usar un procedimiento experimental simple y económicamente eficiente. Cuando uno se encuentra muy lejos del óptimo, por lo general se supone que un modelo de primer orden es una aproximación adecuada de la verdadera superficie en una región pequeña de las x.

El **método del ascenso más pronunciado** es un procedimiento para moverse secuencialmente sobre la trayectoria del ascenso más pronunciado, es decir, en la dirección del incremento máximo en la respuesta. Desde luego, si se desea la **minimización,** entonces se está hablando del **método del descenso más pronunciado**. El modelo de primer orden ajustado es

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i \tag{13-37}$$

y la superficie de respuesta de primer orden (es decir, los contornos de y) es una serie de rectas paralelas como las que se muestran en la figura 13-44. La dirección del ascenso más pronunciado es la dirección en que y se incrementa con mayor rapidez. Esta dirección es normal a los contornos de la superficie de respuesta ajustada. Por lo general se toma como la **trayectoria del ascenso más pronunciado** la recta que pasa por el centro de la región de interés y que es normal a los contornos de la superficie ajustada. Por tanto, los pasos graduales sobre la trayectoria son proporcionales a los coeficientes de regresión $\{\hat{\beta}_i\}$. El experimentador determina el tamaño real de los pasos con base en el conocimiento del proceso o de otras consideraciones de índole práctica.

Se conducen experimentos sobre la trayectoria del ascenso más pronunciado hasta que ya no se observe ningún incremento en la respuesta. Entonces puede ajustarse un nuevo modelo de

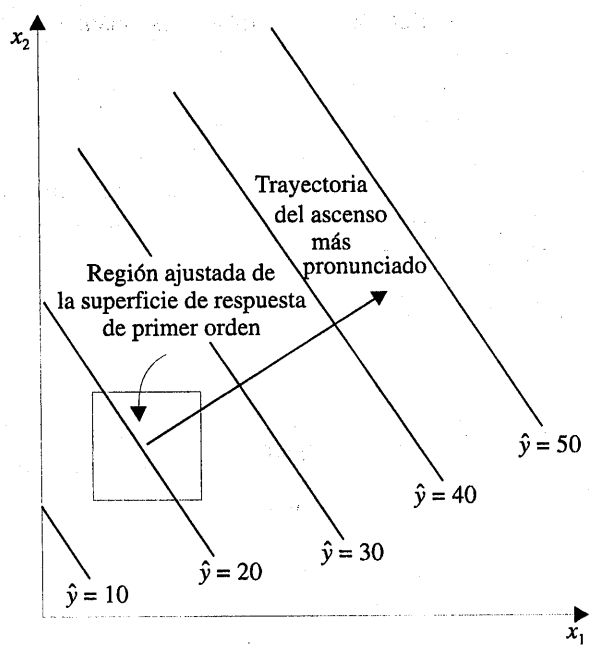


Figura 13-44 Superficie de respuesta de primer orden y trayectoria del ascenso más pronunciado.

primer orden, determinarse una nueva dirección del ascenso más pronunciado, y llevar a cabo experimentos adicionales en esa dirección hasta que el experimentador sienta que el proceso se encuentra cerca del óptimo.

EJEMPLO 13-14

En el ejemplo 13-9 se describió un experimento sobre un proceso químico en el que dos factores, el tiempo de reacción (x_1) y la temperatura de reacción (x_2), afectan la conversión porcentual o rendimiento (Y). En la figura 13-27 se muestra el diseño 2^2 junto con los cinco puntos centrales usados en este estudio. El ingeniero encontró que ambos factores eran importantes, que no había interacción, y que no había curvatura en la superficie de respuesta. Por lo tanto, el modelo de primer orden

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

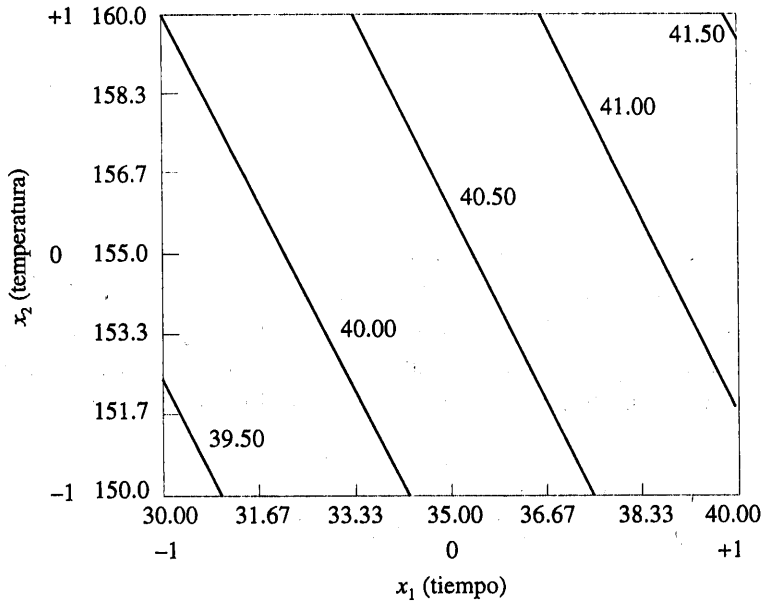
debería ser apropiado. Ahora bien, la estimación del efecto del tiempo es 1.55 y la estimación del efecto de la temperatura es 0.65, y como los coeficientes de regresión $\hat{\beta}_1$ y $\hat{\beta}_2$ son la mitad de las estimaciones de los efectos correspondientes, el modelo de primer orden ajustado es

$$\hat{y} = 40.44 + 0.775x_1 + 0.325x_2$$

En las figuras 13-45*a* y *b* se muestran la gráfica de contorno y la superficie de respuesta tridimensional de este modelo. En la figura 13-45 se muestra asimismo la relación entre las **variables codificadas** x_1 y x_2 (que definieron los niveles alto y bajo de los factores) y las variables originales, el tiempo (en minutos) y la temperatura (en °F).

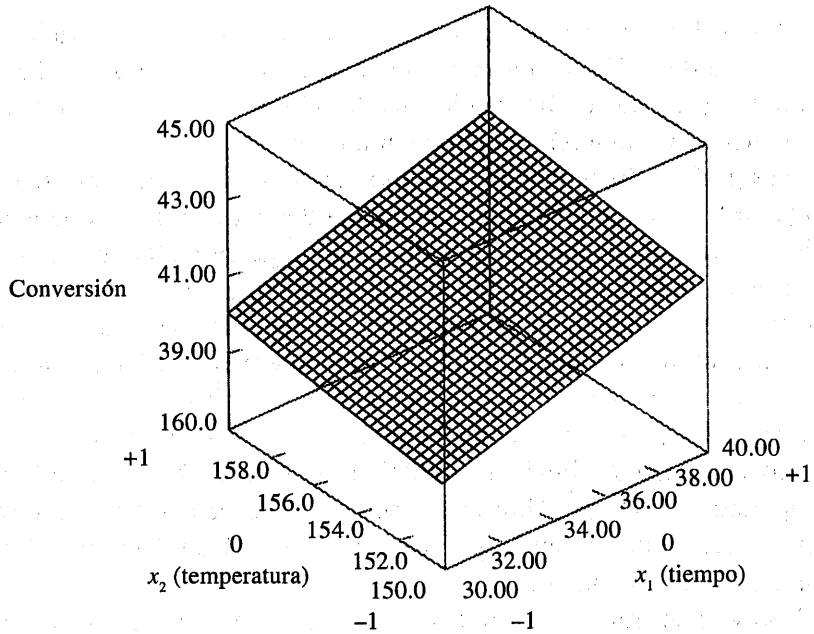
Al examinar estas gráficas (o el modelo ajustado), se observa que al alejarse del centro del diseño —el punto ($x_1 = 0$, $x_2 = 0$)— sobre la trayectoria del ascenso más pronunciado, uno se movería 0.775 unidades en la dirección x_1 por cada 0.325 unidades en la dirección x_2. Por tanto, la trayectoria del ascenso más pronunciado pasa por el punto ($x_1 = 0$, $x_2 = 0$) y tiene pendiente 0.325/0.775. El ingeniero decide usar 5 minutos de tiempo de reacción como el tamaño de cada paso básico. Entonces, 5 minutos de tiempo de reacción es equivalente a un paso en la variable *codificada* x_1 de $\Delta x_1 = 1$. Por lo tanto, los pasos sobre la trayectoria del ascenso más pronunciado son $\Delta x_1 = 1.0000$ y $\Delta x_2 = (0.325/0.775)\ \Delta x_1 = 0.42$. Un cambio de $\Delta x_2 = 0.42$ en la variable codificada x_2 es equivalente a aproximadamente 2°F en la variable original de la temperatura. Por lo tanto, el ingeniero se moverá sobre la trayectoria del ascenso más pronunciado incrementando el tiempo de reacción 5 minutos y la temperatura 2°F. Se determinará una observación real del rendimiento en cada punto.

En la figura 13-46 se muestran varios puntos sobre esta trayectoria del ascenso más pronunciado y los rendimientos observados realmente del proceso en dichos puntos. En los puntos *A-D* el rendimiento observado presenta incrementos uniformes, pero después del punto *D*, el rendimiento disminuye. Por lo tanto, el ascenso más pronunciado terminaría en la vecindad de 55 minutos de tiempo de reacción y 163°F con una conversión porcentual observada de 67%.



Gráfica de contorno

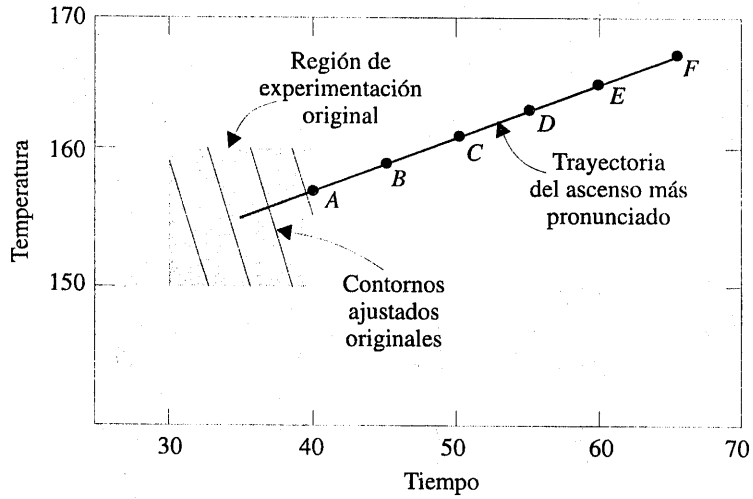
(*a*)



Gráfica de superficie tridimensional

(*b*)

Figura 13-45 Gráficas de superficie de respuesta para el modelo de primer orden del ejemplo 13-14.



Punto A: 40 minutos, 157°F, $y = 40.5$
Punto B: 45 minutos, 159°F, $y = 51.3$
Punto C: 50 minutos, 161°F, $y = 59.6$
Punto D: 55 minutos, 163°F, $y = 67.1$
Punto E: 60 minutos, 165°F, $y = 63.6$
Punto F: 65 minutos, 167°F, $y = 60.7$

Figura 13-46 Experimento del ascenso más pronunciado del ejemplo 13-14.

13-9.2 Análisis de una superficie de respuesta de segundo orden

Cuando el experimentador se encuentra relativamente cerca del óptimo, por lo general se necesita un modelo de segundo orden para aproximar la respuesta, debido a la curvatura de la verdadera superficie de respuesta. El modelo de segundo orden ajustado es

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1}^{k} \hat{\beta}_{ii} x_i^2 + \sum_{i} \sum_{\substack{j \\ i<j}} \hat{\beta}_{ij} x_i x_j$$

donde $\hat{\beta}$ denota la estimación de mínimos cuadrados de β. En esta sección se indicará cómo usar este modelo ajustado para encontrar el conjunto de condiciones de operación óptimas para las x y para caracterizar la naturaleza de la superficie de respuesta.

EJEMPLO 13-15

Continuación del ejemplo 13-14

Considérese el proceso químico del ejemplo 13-14, donde el método del ascenso más pronunciado terminó en un tiempo de reacción de 55 minutos y una temperatura de 162°F. El experimentador decide ajustar un modelo de segundo orden en esta región. En la tabla 13-35 y en la figura 13-47 se muestra el diseño experimental, el cual consta de un diseño 2^2 con centro en 55 minutos y 165°F, de cinco puntos centrales y de cuatro corridas sobre los ejes de coordenadas, llamadas corridas axiales. A este tipo de diseño se le llama **diseño central compuesto**, el cual es muy común para ajustar superficies de respuesta de segundo orden.

Se midieron dos variables de respuesta durante esta fase del experimento: la conversión (rendimiento) porcentual y la viscosidad. El modelo cuadrático de mínimos cuadrados para la variable de respuesta es

$$\hat{y}_1 = 69.1 + 1.633x_1 + 1.083x_2 - 0.969x_1^2 - 1.219x_2^2 + 0.225x_1x_2$$

El análisis de varianza de este modelo se muestra en la tabla 13-36.

Tabla 13-35 Diseño central compuesto para el ejemplo 13-15

Número de observación	Tiempo (minutos)	Temperatura (grados F)	Variables codificadas x_1	x_2	Conversión (por ciento) Respuesta 1	Viscosidad (mPa-seg) Respuesta 2
1	50	160	−1	−1	65.3	35
2	60	160	1	−1	68.2	39
3	50	170	−1	1	66	36
4	60	170	1	1	69.8	43
5	48	165	−1.414	0	64.5	30
6	62	165	1.414	0	69	44
7	55	158	0	−1.414	64	31
8	55	172	0	1.414	68.5	45
9	55	165	0	0	68.9	37
10	55	165	0	0	69.7	34
11	55	165	0	0	68.5	35
12	55	165	0	0	69.4	36
13	55	165	0	0	69	37

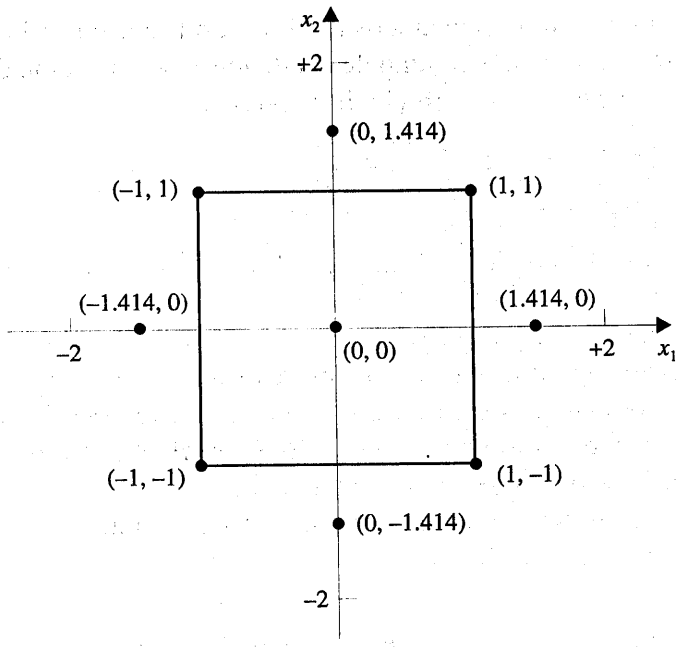


Figura 13-47 Diseño central compuesto para el ejemplo 13-15.

Tabla 13-36 Análisis de varianza del modelo cuadrático, respuesta rendimiento

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Modelo	45.89	5	9.178	14.93	0.0013
Residual	4.30	7	0.615		
Total	50.19	12			

Variable independiente	Estimación de coeficientes	Error estándar	t para H_0 Coeficiente = 0	Valor P
Ordenada al origen	69.100	0.351	197.1	
x_1	1.633	0.277	5.891	0.0006
x_2	1.083	0.277	3.907	0.0058
x_1^2	–0.969	0.297	–3.259	0.0139
x_2^2	–1.219	0.297	–4.100	0.0046
x_1x_2	0.225	0.392	0.5740	0.5839

En la figura 13-48 se muestra la gráfica de contorno de la superficie de respuesta y la gráfica de superficie tridimensional de este modelo. Al examinar estas gráficas, el rendimiento máximo es de cerca de 70%, obtenido con aproximadamente 60 minutos de tiempo de reacción y 167°F.

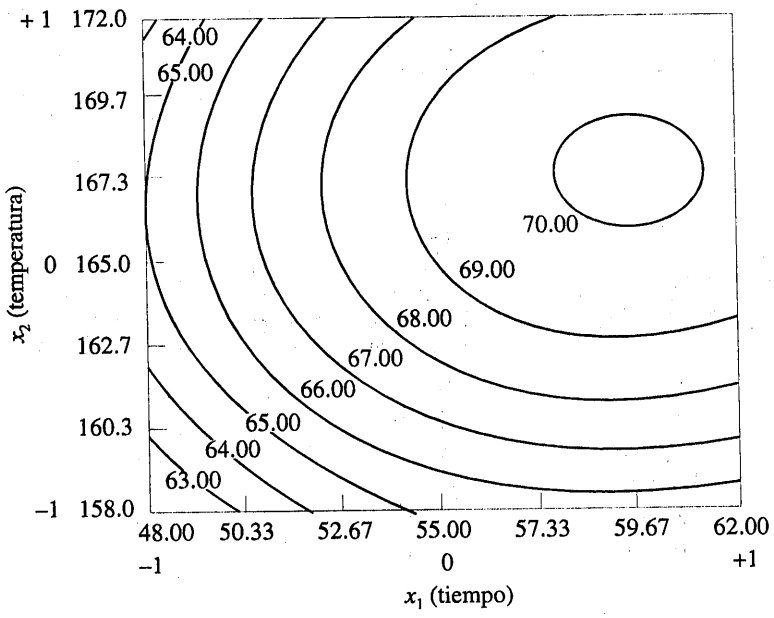
La respuesta viscosidad la describe adecuadamente el modelo de primer orden

$$\hat{y}_2 = 37.08 + 3.85x_1 + 3.10x_2$$

En la tabla 13-37 se resume el análisis de varianza de este modelo. La superficie de respuesta se ilustra gráficamente en la figura 13-49. Obsérvese que la viscosidad se incrementa cuando tanto el tiempo como la temperatura se incrementan.

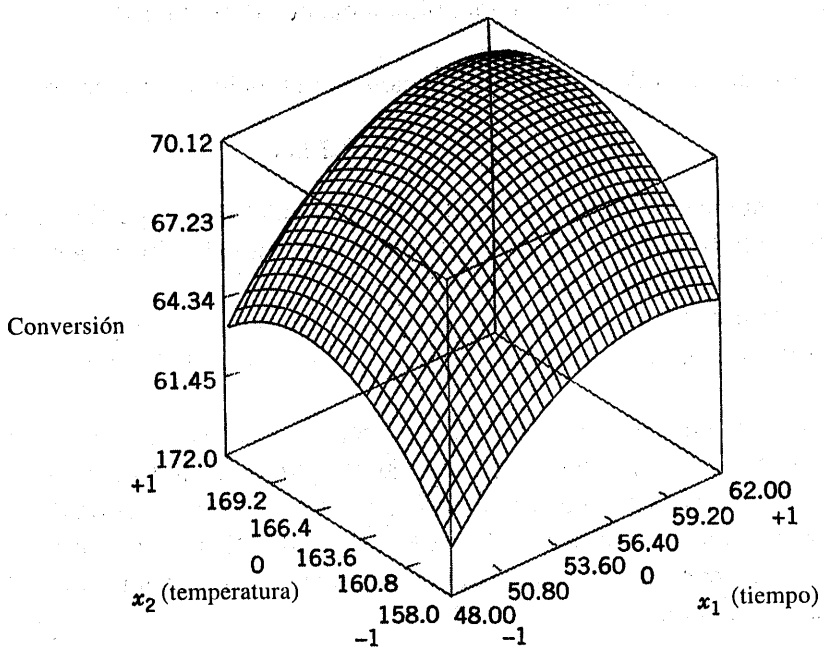
Como en la mayoría de los problemas de superficie de respuesta, el experimentador de este ejemplo tenía objetivos antagónicos en cuanto a las dos respuestas. El objetivo era maximizar el rendimiento, pero el rango aceptable para la viscosidad era $38 \leq y_2 \leq 42$. Cuando sólo hay un número reducido de variables independientes, una forma sencilla de resolver este problema es superponer las superficies de respuesta para encontrar el óptimo. En la figura 13-50 se muestra la gráfica superpuesta de ambas respuestas, con los contornos y_1 = conversión de 69%, $y_2 = 38$ y $y_2 = 42$ resaltados. Las áreas sombreadas de esta gráfica identifican combinaciones impracticables del tiempo y la temperatura. Esta gráfica muestra que varias combinaciones del tiempo y la temperatura serán satisfactorias.

En el ejemplo 13-15 se ilustra el uso de un **diseño central compuesto** (DCC) para ajustar un modelo de superficie de respuesta de segundo orden. Estos diseños son de uso generalizado en la práctica porque son relativamente eficientes en lo que se refiere al número de corridas requeridas. En general, un DCC con k factores requiere 2^k corridas factoriales, $2k$ corridas axiales



Gráfica de superficie

(*a*)



Gráfica de superficie

(*b*)

Figura 13-48 Gráficas de superficie de respuesta en la respuesta del rendimiento, ejemplo 13-15.

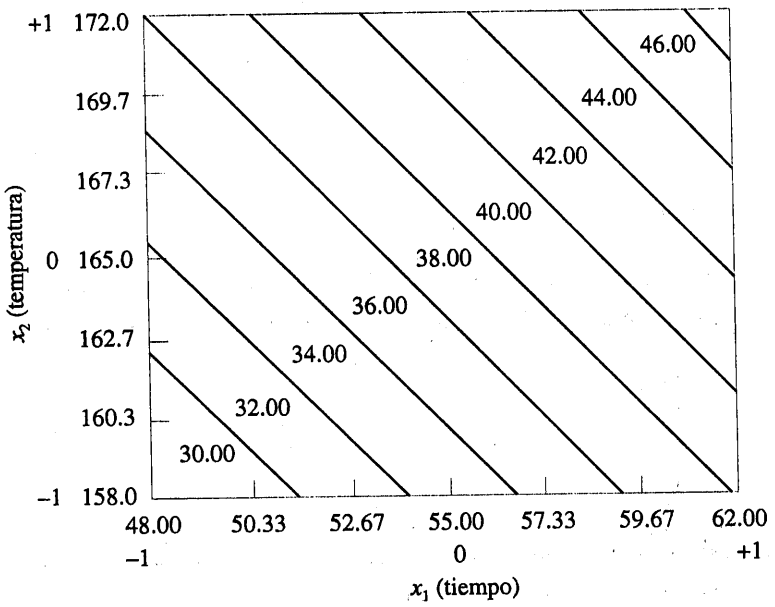
Tabla 13-37 Análisis de varianza para el modelo de primer orden, respuesta viscosidad

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Modelo	195.4	2	97.728	15.89	0.0008
Residual	61.5	10	6.15		
Total	256.9	12			

Variable independiente	Estimación del coeficiente	Grados de libertad	Error estándar	t para H_0 Coeficiente = 0	Valor P
Ordenada al origen	37.08	1	0.69	53.91	
x_1	3.85	1	0.88	4.391	0.0014
x_2	3.10	1	0.88	3.536	0.0054

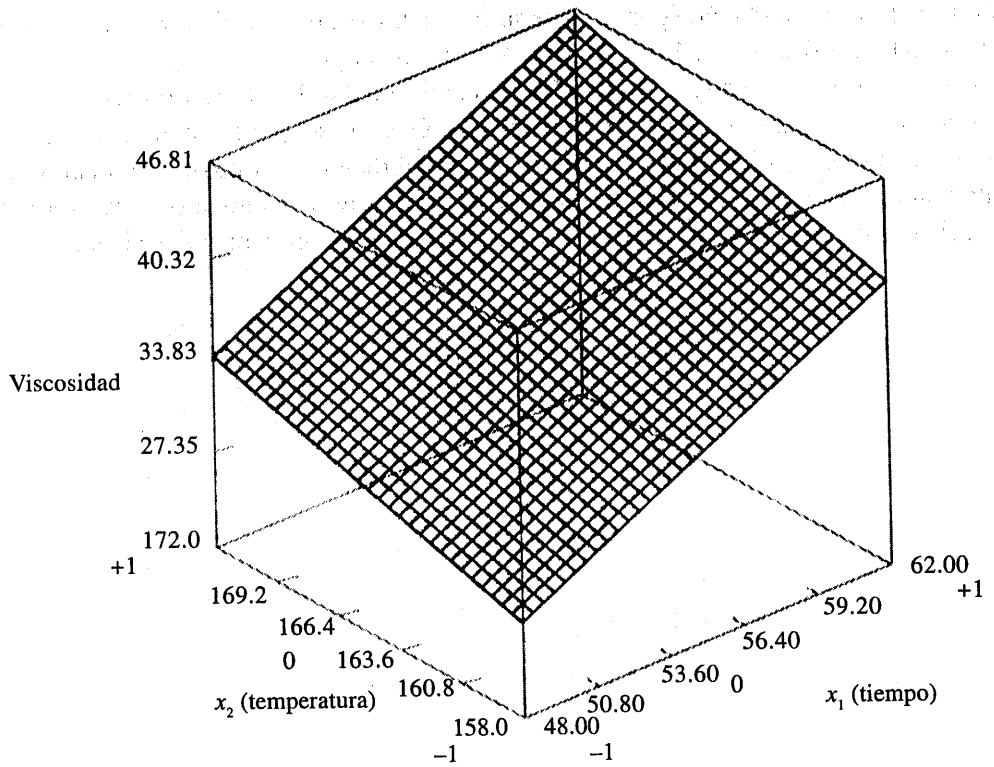
y al menos un punto central (lo más común es usar entre 3 y 5 puntos centrales). En la figura 13-51 se muestran los diseños para $k = 2$ y $k = 3$ factores.

El diseño central compuesto puede hacerse **girable** mediante la elección adecuada de la separación axial α en la figura 13-51. Si el diseño es girable, la desviación estándar de la respuesta predicha $\hat{y}$ es constante en todos los puntos que están a la misma distancia desde el centro del diseño. Para hacer el diseño girable se elige $\alpha = (F)^{1/4}$, donde F es el número de puntos en la parte factorial del diseño (por lo general $F = 2^k$). Para el caso de $k = 2$ factores, $\alpha = (2^2)^{1/4} = 1.414$, como se usó en el diseño del ejemplo 13-15. En la figura 13-52 se presenta una gráfica de contorno y una gráfica de superficie de la desviación estándar de la predicción para el modelo cuadrático usado para la respuesta rendimiento. Obsérvese que los contornos son círculos concéntricos, lo cual implica que el rendimiento se predice con la misma precisión para todos los puntos que están a la misma distancia del centro del diseño. Asimismo, como sería de esperarse, la precisión disminuye con el incremento de la distancia del centro del diseño.



Gráfica de contorno

(*a*)



Gráfica de superficie

(*b*)

Figura 13-49 Gráficas de superficie de respuesta para la respuesta viscosidad, ejemplo 13-15.

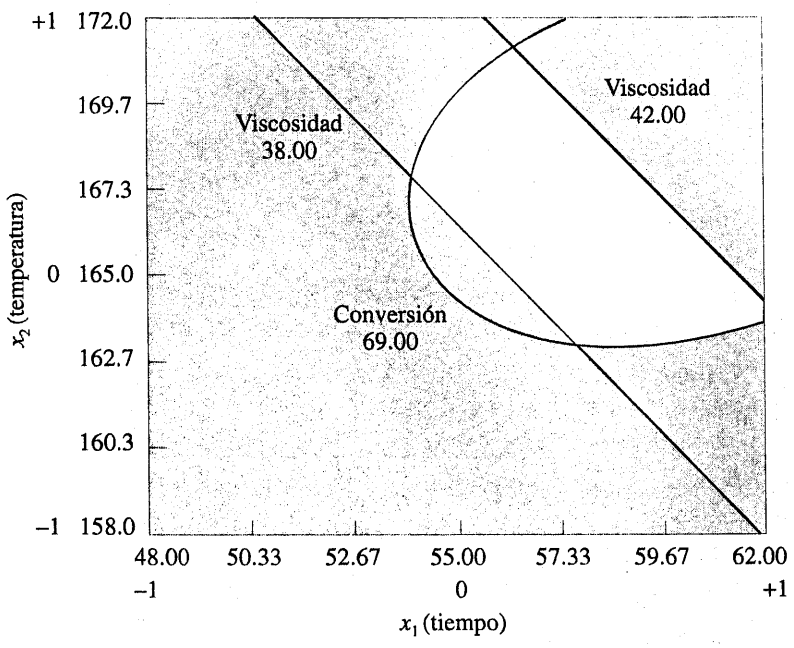


Figura 13-50 Suposición de las superficies de respuesta del rendimiento y la viscosidad, ejemplo 13-15.

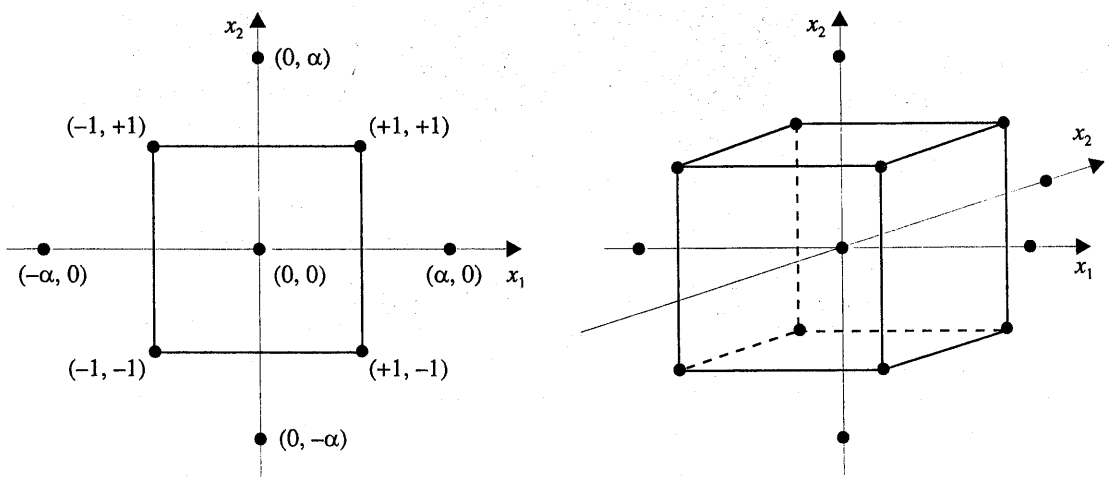
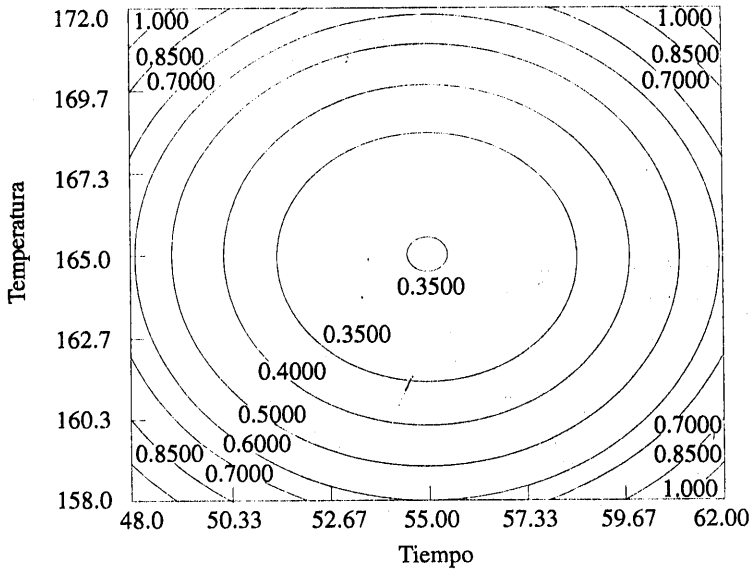
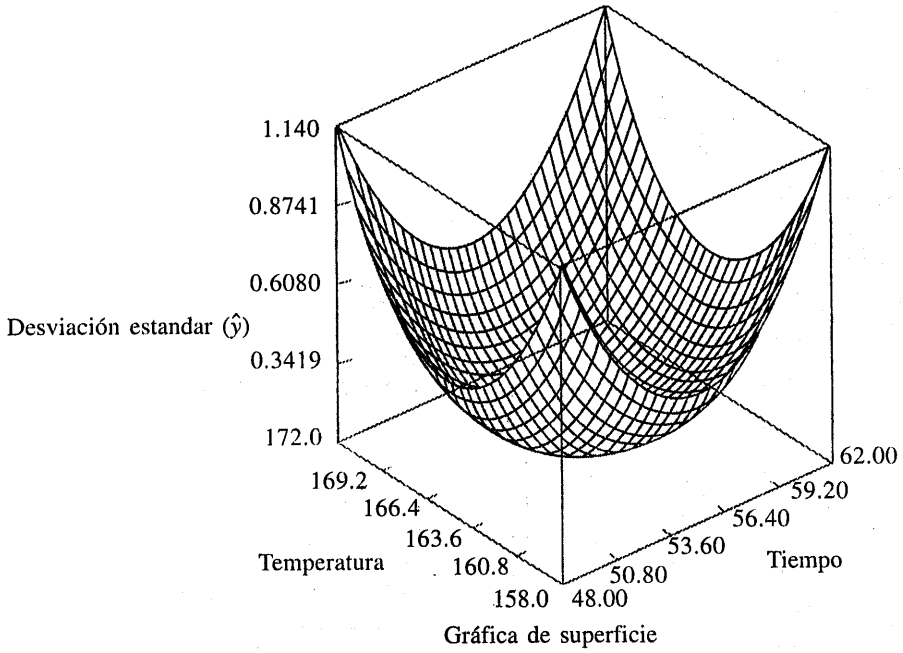


Figura 13-51 Diseños centrales compuestos para $k = 2$ y $k = 3$.



Gráfica de contorno

(*a*)



Gráfica de superficie

(*b*)

Figura 13-52 Gráficas de $\sqrt{V(\hat{y})}$ constante para un diseño central compuesto girable.

EJERCICIOS DE LA SECCIÓN 13-9

13-41. En un artículo de *Rubber Age* (Vol. 89, pp. 453-458) se describe un experimento sobre la fabricación de un producto en el que se hicieron variar dos factores. Los factores son el tiempo de reacción (hr) y la temperatura (°C). Estos factores se codifican como x_1 = (tiempo – 12)/8 y x_2 = (temperatura – 250)/30. Se observaron los datos siguientes, donde y es el rendimiento (en por ciento)

Número de corrida	x_1	x_2	y
1	–1	0	83.8
2	1	0	81.7
3	0	0	82.4
4	0	0	82.9
5	0	–1	84.7
6	0	1	75.9
7	0	0	81.2
8	–1.414	–1.414	81.3
9	–1.414	1.414	83.1
10	1.414	–1.414	85.3
11	1.414	1.414	72.7
12	0	0	82.0

a) Grafique los puntos en los que se hicieron las corridas experimentales.
b) Ajuste un modelo de segundo orden a los datos. ¿El modelo de segundo orden es adecuado?
c) Trace la superficie de respuesta del rendimiento. ¿Qué recomendaciones se harían sobre las condiciones de operación de este proceso?

13-42. Considérese el modelo de primer orden

$$\hat{y} = 50 + 1.5x_1 - 0.8x_2$$

donde $-1 \le x_i \le 1$. Encuentre la dirección del ascenso más pronunciado.

13-43. Un fabricante de herramientas de corte ha desarrollado dos ecuaciones empíricas para la vida de la herramienta (y_1) y el costo de la herramienta (y_2). Ambos modelos son funciones de la dureza de la herramienta (x_1) y del tiempo de fabricación (x_2). Las ecuaciones son

$$\hat{y}_1 = 10 + 5x_1 + 2x_2$$
$$\hat{y}_2 = 23 + 3x_1 + 4x_2$$

y ambas ecuaciones son válidas en el rango $-1.5 \le x_i \le 1.5$. Suponga que la vida de la herramienta debe exceder 12 horas y que el costo debe ser menor de $27.50.

a) ¿Existe un conjunto factible de condiciones de operación?
b) ¿En dónde debería realizarse este proceso?

13-44. En un artículo de *Tappi* (Vol. 43, pp. 38-44) se describe un experimento para investigar el valor en ceniza (una medición de las impurezas inorgánicas) de la pulpa de papel. Se estudiaron dos variables, la temperatura t en grados Celsius y el tiempo t en horas, y parte de los resultados se muestran en la siguiente tabla. Las variables predictoras codificadas son

$$x_1 = \frac{(T - 775)}{115}, \quad x_2 = \frac{(t - 3)}{1.5}$$

y la respuesta y es (valor en ceniza en seco en %) $\times 10^3$.

x_1	x_2	y
–1	–1	211
1	–1	92
–1	1	216
1	1	99
–1.5	0	222
1.5	0	48
0	–1.5	168
0	1.5	179
0	0	122
0	0	175
0	0	157
0	0	146

a) ¿Qué tipo de diseño se ha usado en este estudio? ¿El diseño es girable?
b) Ajuste un modelo cuadrático a los datos. ¿Este modelo es satisfactorio?
c) Si es importante minimizar el valor en ceniza, ¿en dónde debería correrse el proceso?

13-45. En el libro *Empirical Model Building and Response Surfaces* (John Wiley), G.E.P. Box y N.R. Draper describen un experimento con tres factores. Los datos que se muestran en la tabla siguiente son una variación del experimento original de la página 247 de su libro. Suponga que estos datos se recabaron en un proceso de manufactura de semiconductores.

x_1	x_2	x_3	y_1	y_2
−1	−1	−1	24.00	12.49
0	−1	−1	120.33	8.39
1	−1	−1	213.67	42.83
−1	0	−1	86.00	3.46
0	0	−1	136.63	80.41
1	0	−1	340.67	16.17
−1	1	−1	112.33	27.57
0	1	−1	256.33	4.62
1	1	−1	271.67	23.63
−1	−1	0	81.00	0.00
0	−1	0	101.67	17.67
1	−1	0	357.00	32.91
−1	0	0	171.33	15.01
0	0	0	372.00	0.00
1	0	0	501.67	92.50
−1	1	0	264.00	63.50
0	1	0	427.00	88.61
1	1	0	730.67	21.08
−1	−1	1	220.67	133.82
0	−1	1	239.67	23.46
1	−1	1	422.00	18.52
−1	0	1	199.00	29.44
0	0	1	485.33	44.67
1	0	1	673.67	158.21
−1	1	1	176.67	55.51
0	1	1	501.00	138.94
1	1	1	1 010.00	142.45

a) La respuesta y_1 es el promedio de tres lecturas de resistividad de una oblea individual. Ajuste un modelo lineal a esta respuesta.

b) La respuesta y_2 es la desviación estándar de las tres mediciones de resistividad. Ajuste un modelo cuadrático a esta respuesta.

c) ¿En dónde se recomendaría establecer x_1, x_2 y x_3 si el objetivo es mantener la resistividad media en 500 y minimizar la desviación estándar?

Ejercicios complementarios

13-46. En un artículo de *Process Engineering* (No. 71, pp. 46-47) se presenta un experimento factorial con dos factores usado para investigar el efecto del pH y la concentración del catalizador en la viscosidad de un producto (cSt). Los datos son los siguientes:

		Concentración del catalizador	
		2.5	2.7
pH	5.6	192, 199, 189, 198	178, 186, 179, 188
	5.9	185, 193, 185, 192	197, 196, 204, 204

a) Pruebe los efectos principales y las interacciones usando $\alpha = 0.05$. ¿Cuáles son las conclusiones?

b) Grafique la interacción y comente la información proporcionada por esta gráfica.

c) Analice los residuales de este experimento.

13-47. El tratamiento térmico de piezas metálicas es un procedimiento común en procesos de manufactura. En un artículo de *Journal of Metals* (Vol. 41) se describe un experimento para investigar la distorsión del aplanado producida por el tratamiento térmico para tres tipos de engranes y dos tiempos de tratamiento térmico. Los datos son los siguientes:

	Tiempo (minutos)	
Tipo de engrane	90	120
20-dientes	0.0265	0.0560
	0.0340	0.0650
24-dientes	0.0430	0.0720
	0.0510	0.0880
28-dientes	0.0405	0.0620
	0.0575	0.0825

a) ¿Hay evidencia de que la distorsión del aplanado es superficial para los diferentes tipos de engranes? ¿Hay algún indicio de que el tiempo de tratamiento térmico afecta la distorsión del aplanado? ¿Hay interacción entre estos factores? Use $\alpha = 0.05$.

b) Construya las gráficas de los efectos de los factores que ayuden a sacar conclusiones de este experimento.

c) Analice los residuales de este experimento. Comente la validez de los supuestos fundamentales.

13-48. En un artículo de *Textile Research Institute Journal* (Vol. 54, pp. 171-179) se reportan los resultados de un experimento que estudiaba los efectos de tratar un género con sales inorgánicas seleccionadas y la combustibilidad del material. Se usaron dos niveles de aplicación de cada sal, y se usó una prueba de quemado vertical en cada muestra. (Con esta prueba se encuentra la temperatura a la que se enciende cada muestra.) Los datos de prueba de quemado se indican abajo.

a) Pruebe las diferencias entre las sales, los niveles de aplicación y las interacciones. Use $\alpha = 0.01$.

b) Trace una gráfica de la interacción entre la sal y el nivel de aplicación. ¿Qué conclusiones pueden sacarse de esta gráfica?

c) Analice los residuales de este experimento.

13-49. En un artículo de *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* (Vol. 15) se describe un experimento para investigar un método para alinear chips ópticos en tarjetas de circuitos. El método incluye la colocación circunferencial de soldadura en la parte inferior del chip. En el experimento se usaron tres tamaños de soldadura y tres métodos de alineación. La variable de respuesta es la precisión de la alineación (μm). Los datos son los siguientes:

Tamaño de la soldadura de colocación circunferencial (diámetro en μm)	Método de alineación 1	2	3
	4.60	1.55	1.05
75	4.53	1.45	1.00
	2.33	1.72	0.82
130	2.44	1.76	0.95
	4.95	2.73	2.36
260	4.55	2.60	2.46

a) ¿Hay algún indicio de que el tamaño de soldadura o el método de alineación afecta la precisión de la alineación? ¿Hay evidencia de interacción entre estos factores? Use $\alpha = 0.05$.

b) ¿Qué recomendaciones se harían acerca de este proceso?

c) Analice los residuales de este experimento. Comente la adecuación del modelo.

Nivel	Sal: Sin tratamiento	$MgCl_2$	NaC_1	$CaCO_3$	$CaCl_2$	Na_2CO_3
1	812	752	739	733	725	751
	827	728	731	728	727	761
	876	764	726	720	719	755
2	945	794	741	786	756	910
	881	760	744	771	781	854
	919	757	727	779	814	848

13-50. En un artículo de *Solid State Technology* (Vol. 29, pp. 281-284) se describe el uso de experimentos factoriales en fotolitografía, un paso importante en el proceso de fabricación de circuitos integrados. La variables de este experimento (todas con dos niveles) son la temperatura de precocido (A), tiempo de precocido (B) y la energía de exposición (C), y la variable de respuesta es la anchura de la línea delta, la diferencia entre la línea de la plantilla y la línea impresa en el dispositivo. Los datos son los siguientes: $(1) = -2.30$, $a = -9.87$, $b = -18.20$, $ab = -30.20$, $c = -23.80$, $ac = -4.30$, $bc = -3.80$, y $abc = -14.70$.

a) Estime los efectos de los factores.

b) Suponga que se agrega un punto central a este diseño y se obtienen cuatro réplicas: –10.50, –5.30, –11.60 y –7.30. Calcule una estimación del error experimental.

c) Pruebe la significación de los efectos principales, las interacciones y la falta de ajuste. Con $\alpha = 0.05$, ¿qué conclusiones pueden sacarse?

d) ¿Qué modelo se recomendaría para predecir la respuesta de la anchura de la línea delta, con base en los resultados de este experimento?

e) Analice los residuales de este experimento y comente la adecuación del modelo.

13-51. En un artículo de *Journal of Coatings Technology* (Vol. 60, pp. 27-32) se describe un diseño factorial 2^4 usado para estudiar el recubrimiento de plata usado en automóviles. La variable de respuesta es la definición de la imagen (DDI). Las variables usadas en el experimento son:

$A =$ Por ciento de poliéster por peso de poliéster/melamina (valor bajo = 50%, valor alto = 70%)

$B =$ Por ciento de carboxilato butirato acetato de celulosa (valor bajo = 15%, valor alto = 30%)

$C =$ Por ciento de estearato de aluminio (valor bajo = 1%, valor alto = 3%)

$D =$ Por ciento de catalizador ácido (valor bajo = 0.25%, valor alto = 0.50%)

Las respuestas son $(1) = 63.8$, $a = 77.6$, $b = 68.8$, $ab = 76.5$, $c = 72.5$, $ac = 77.2$, $bc = 77.7$, $abc = 84.5$, $d = 60.6$, $ad = 64.9$, $bd = 72.7$, $abd = 73.3$, $cd = 68.0$, $acd = 76.3$, $bcd = 76.0$ y $abcd = 75.9$.

a) Estime los efectos de los factores.

b) Usando una gráfica de probabilidad normal de los efectos, identifique un modelo tentativo para los datos de este experimento.

c) Usando los factores aparentemente insignificantes como estimación del error, pruebe la significación de los factores identificados en el inciso b). Use $\alpha = 0.05$.

d) ¿Qué modelo se usaría para describir el proceso, con base en este experimento? Interprete el modelo.

e) Analice los residuales del modelo del inciso d) y comente los resultados.

13-52. En un artículo de *Journal of Manufacturing Systems* (Vol. 10, pp. 32-40) se describe un experimento para investigar el efecto de cuatro factores —P = presión del chorro, F = índice de flujo abrasivo, G = tamaño del grano abrasivo y V = velocidad transversal del chorro— sobre la rugosidad superficial de un cortador de chorro de agua. En la página siguiente se muestra un diseño 2^4 con siete puntos centrales.

a) Estime los efectos de los factores.

b) Construya un modelo tentativo examinando una gráfica de probabilidad normal de los efectos.

c) ¿El modelo del inciso b) es una descripción razonable del proceso? ¿Es significativa la falta de ajuste? Use $\alpha = 0.05$.

d) Interprete los resultados de este experimento.

e) Analice los residuales de este experimento.

13-53. Construya un diseño 2_{IV}^{4-1} para el problema del ejercicio 13-51. Seleccione los datos para las ocho corridas que se habrían requerido en este diseño. Analice estas corridas y compare las conclusiones con las

Corrida	Factores V (pulg/min)	F (lb/min)	P kpsi	G (Malla No.)	Rugosidad superficial (μm)
1	6	2.0	38	80	104
2	2	2.0	38	80	98
3	6	2.0	30	80	103
4	2	2.0	30	80	96
5	6	1.0	38	80	137
6	2	1.0	38	80	112
7	6	1.0	30	80	143
8	2	1.0	30	80	129
9	6	2.0	38	170	88
10	2	2.0	38	170	70
11	6	2.0	30	170	110
12	2	2.0	30	170	110
13	6	1.0	38	170	102
14	2	1.0	38	170	76
15	6	1.0	30	170	98
16	2	1.0	30	170	68
17	4	1.5	34	115	95
18	4	1.5	34	115	98
19	4	1.5	34	115	100
20	4	1.5	34	115	97
21	4	1.5	34	115	94
22	4	1.5	34	115	93
23	4	1.5	34	115	91

que se obtuvieron en el ejercicio 13-51 para el diseño factorial completo.

13-54. Construya un diseño 2_{IV}^{4-1} para el problema del ejercicio 13-52. Seleccione los datos para las ocho corridas que se habrían requerido en este diseño, además de los puntos tos centrales. Analice estos datos y compare las conclusiones con las que se obtuvieron en el ejercicio 13-52 para el diseño factorial completo.

13-55. Construya un diseño 2_{IV}^{8-4} con 16 corridas. ¿Cuáles son las relaciones de los alias en este diseño?

13-56. Construya un diseño 2_{III}^{5-2} con 8 corridas. ¿Cuáles son las relaciones de los alias en este diseño?

13-57. En un estudio sobre el rendimiento en el desarrollo de un proceso, se estudiaron cuatro factores, cada uno con dos niveles: el tiempo (A), la concentración (B), la presión (C) y la temperatura (D). Se corre una sola réplica de un diseño 2^4, y los datos resultantes se presentan en la tabla el principio de la página siguiente.

a) Grafique las estimaciones de los efectos en una escala de probabilidad normal. ¿Qué factores parecen tener efectos grandes?

b) Realice un análisis de varianza usando la gráfica de probabilidad normal del inciso a) como guía para formar el término del error. ¿A qué conclusiones se llega?

c) Analice los residuales de este experimento. ¿El análisis indica la presencia de algún problema potencial?

d) ¿Este diseño puede reducirse a un diseño 2^3 con dos réplicas? De ser así, trace el diseño anotando el promedio y el rango del rendimiento en cada punto del cubo. Interprete los resultados.

13-58. Considérese el experimento descrito en el ejercicio 13-57. Encuentre intervalos de confianza de 95% para los efectos de los factores que parecen ser importantes. Use la gráfica de probabilidad normal como guía para identificar los efectos que pueden combinarse para proporcionar una estimación del error.

13-59. En un artículo de *Journal of Quality Technology* (Vol. 17, pp. 198-206) se describe el uso de un diseño factorial fraccionado con réplicas para investigar el efecto de cinco factores sobre la altura libre de los resortes de hojas usados en una aplicación automotriz. Los factores son A = temperatura del horno, B = tiempo de calentamiento, C = tiempo de transferencia, D = tiempo de retención, y E = temperatura del aceite de revenido. Los datos se muestran en la siguiente tabla.

a) ¿Cuál es el generador de esta fracción? Escriba la estructura de los alias.

b) Analice los datos. ¿Qué factores influyen en la altura libre promedio?

c) Calcule el rango de la altura libre de cada corrida. ¿Hay algún indicio de que cualquiera de estos factores afecta la variabilidad de la altura libre?

Número de corrida	Orden real de las corridas	*A*	*B*	*C*	*D*	Rendimiento (libras)
1	5	−	−	−	−	12
2	9	+	−	−	−	18
3	8	−	+	−	−	13
4	13	+	+	−	−	16
5	3	−	−	+	−	17
6	7	+	−	+	−	15
7	14	−	+	+	−	20
8	1	+	+	+	−	15
9	6	−	−	−	+	10
10	11	+	−	−	+	25
11	2	−	+	−	+	13
12	15	+	+	−	+	24
13	4	−	−	+	+	19
14	16	+	−	+	+	21
15	10	−	+	+	+	17
16	12	+	+	+	+	23

Niveles de los factores	Bajo (−)	Alto (+)
A (h)	2.5	3
B (%)	14	18
C (psi)	60	80
D (°C)	225	250

d) Analice los residuales de este experimento y comente los resultados.

A	*B*	*C*	*D*	*E*	Altura libre		
−	−	−	−	−	7.78	7.78	7.81
+	−	−	+	−	8.15	8.18	7.88
−	+	−	+	−	7.50	7.56	7.50
+	+	−	−	−	7.59	7.56	7.75
−	−	+	+	−	7.54	8.00	7.88
+	−	+	−	−	7.69	8.09	8.06
−	+	+	−	−	7.56	7.52	7.44
+	+	+	+	−	7.56	7.81	7.69
−	−	−	−	+	7.50	7.56	7.50
+	−	−	+	+	7.88	7.88	7.44
−	+	−	+	+	7.50	7.56	7.50
+	+	−	−	+	7.63	7.75	7.56
−	−	+	+	+	7.32	7.44	7.44
+	−	+	−	+	7.56	7.69	7.62
−	+	+	−	+	7.18	7.18	7.25
+	+	+	+	+	7.81	7.50	7.59

13-60. En un artículo de *Rubber Chemistry and Technology* (Vol. 47, pp. 825-836) se describe un experimento que estudia la viscosidad Mooney del hule con varias variables, incluyendo el agregado de sílice (partes por centena) y el agregado de aceite (partes por centena). A continuación se presentan parte de los datos de este experimento, donde

$$x_1 = \frac{\text{sílice} - 60}{15}, \quad x_2 = \frac{\text{aceite} - 21}{15}$$

Niveles codificados x_1	x_2	y
−1	−1	13.71
1	−1	14.15
−1	1	12.87
1	1	13.53
−1.4	0	12.99
1.4	0	13.89
0	−1.4	14.16
0	1.4	12.90
0	0	13.75
0	0	13.66
0	0	13.86
0	0	13.63
0	0	13.74

a) ¿Qué tipo de diseño experimental se ha usado? ¿Es girable?

b) Ajuste un modelo cuadrático a estos datos. ¿Qué valores de x_1 y x_2 maximizarán la viscosidad Mooney?

EJERCICIOS PARA DESARROLLAR EL INTELECTO

13-61. Considérese un diseño factorial 2^k no replicado, y suponer que falta una de las combinaciones de tratamientos. Un enfoque lógico de este problema es estimar el valor faltante mediante un número que haga cero la interacción de orden más alto. Aplique esta técnica a los datos del ejemplo 13-8, suponiendo que falta *ab*. Compare los resultados del análisis de estos datos con los resultados del ejemplo 13-8.

13-62. ¿Qué esquema de descomposición en bloques se recomendaría si fuera necesario correr un diseño 2^4 en cuatro bloques de cuatro corridas cada uno?

13-63. Considérese un diseño 2^2 en dos bloques, con *AB* confundidos con los bloques. Demuestre algebraicamente que $SS_{AB} = SS_{\text{Bloques}}$.

13-64. Considérese un diseño 2^3. Suponga que el número mayor de corridas que pueden procesarse en un bloque es cuatro, pero que pueden realizarse un total de 32 observaciones.

a) Sugiera un esquema de descomposición en bloques que proporcione cierta información sobre todas las interacciones.

b) Muestre un plan general (sólo fuente de variabilidad y grados de libertad) para el análisis de varianza de este diseño.

13-65. Construya un diseño 2^{5-1}. Suponga que este diseño debe realizarse en dos bloques de ocho corridas cada uno. Indique cómo puede hacerse esto confundiendo una interacción de dos factores (y el alias de su interacción de tres factores) con los bloques.

13-66. Construya un diseño 2_{IV}^{7-2}. Indique cómo puede confundirse este diseño en cuatro bloques de ocho corridas cada uno. ¿Alguna de las interacciones de dos factores está confundida con los bloques?

13-67. Construya un diseño 2_{IV}^{7-3}. Indique cómo puede confundirse este diseño en dos bloques de ocho corridas cada uno sin perder información sobre cualquier interacción de dos factores.

13-68. Establezca un diseño 2_{III}^{7-4} usando $D = AB$, $E = AC$, $F = BC$ y $G = ABC$ como generadores del diseño. Ignore todas las interacciones con más de dos factores.

a) Verifique que cada efecto principal es alias de las tres interacciones de dos factores.

b) Suponga que se realiza un segundo diseño 2_{III}^{7-4} con generadores $D = -AB$, $E = -AC$, $F = -BC$ y $G = ABC$. ¿Cuáles son los alias de los efectos principales en este diseño?

c) ¿Qué factores pueden estimarse si se combinan los dos conjuntos anteriores de las estimaciones de los efectos de los factores?

14 Estadística no paramétrica

PLAN GENERAL DEL CAPÍTULO

14-1 INTRODUCCIÓN

14-2 PRUEBA DE SIGNOS

- 14-2.1 Descripción de la prueba
- 14-2.2 Prueba de signos para muestras pareadas
- 14-2.3 Error tipo II para la prueba de signos
- 14-2.4 Comparación con la prueba *t*

14-3 PRUEBA DE RANGOS CON SIGNO DE WILCOXON

- 14-3.1 Descripción de la prueba
- 14-3.2 Aproximación para muestras grandes
- 14-3.3 Observaciones pareadas
- 14-3.4 Comparación con la prueba *t*

14-4 PRUEBA DE LA SUMA DE RANGOS DE WILCOXON

- 14-4.1 Descripción de la prueba
- 14-4.2 Aproximación con muestras grandes
- 14-4.3 Comparación con la prueba *t*

14-5 MÉTODOS NO PARAMÉTRICOS EN EL ANÁLISIS DE VARIANZA

- 14-5.1 Prueba de Kruskal-Wallis
- 14-5.2 Transformación de rangos

14-1 INTRODUCCIÓN

La mayoría de los procedimientos para probar hipótesis y construir intervalos de confianza que se han revisado en los capítulos precedentes parten del supuesto de que se está trabajando con muestras aleatorias de poblaciones normales. Tradicionalmente, a estos procedimientos se les ha llamado **métodos paramétricos** porque están basados en una familia paramétrica particular

de distribuciones; en este caso, la distribución normal. De manera alternativa, se dice que estos procedimientos no son *de distribución libre* porque dependen del supuesto de normalidad. Por fortuna, la mayoría de estos procedimientos son relativamente insensibles a desviaciones ligeras de la normalidad. En general, las pruebas t y F y los intervalos de confianza t tendrán niveles de significación o niveles de confianza reales que difieren de los niveles nominales o anunciados que haya elegido el experimentador, si bien la diferencia entre los niveles reales y los anunciados por lo general es bastante reducida cuando la población fundamental no es muy diferente de la normal.

En el presente capítulo se describen los procedimientos llamados **métodos no paramétricos** y **métodos de distribución libre** y, por lo general, el único supuesto que se establece respecto de la distribución de la población fundamental es que sea continua. Estos procedimientos tienen niveles reales de significación α o niveles de confianza del $100(1 - \alpha)$ por ciento para muchos tipos diferentes de distribuciones. Estos procedimientos poseen interesantes atractivos. Una de sus ventajas es que no es necesario que los datos sean cualitativos, sino que pueden ser datos categóricos (como sí o no, defectuoso o no defectuoso) o datos en rangos. Otra ventaja es que por lo general la aplicación de los procedimientos no paramétricos es muy rápida y sencilla.

Los procedimientos que se describen en este capítulo compiten con los procedimientos paramétricos t y F descritos anteriormente. Por consiguiente, es importante comparar el desempeño de los métodos paramétricos y los no paramétricos bajo los supuestos de poblaciones normales y no normales. En general, los procedimientos no paramétricos no utilizan toda la información que proporciona una muestra. Como resultado, la eficiencia de un procedimiento no paramétrico será menor que la del procedimiento paramétrico correspondiente cuando la población fundamental es normal. Esta pérdida de eficiencia se refleja en el requisito de un tamaño de la muestra más grande en un procedimiento no paramétrico que el que se requeriría en un procedimiento paramétrico para alcanzar la misma potencia. Por otra parte, en general esta pérdida de eficiencia no es grande, y con frecuencia la diferencia en el tamaño de las muestras es pequeña. Cuando las distribuciones fundamentales no están cerca de una distribución normal, los métodos no paramétricos tienen mucho que ofrecer. Con frecuencia proporcionan una mejora considerable en comparación con los métodos paramétricos de la teoría de distribución normal.

En general, si es posible aplicar tanto métodos paramétricos como no paramétricos en un problema particular, se usaría el procedimiento paramétrico más eficiente. Sin embargo, los supuestos del método paramétrico pueden ser difíciles o imposibles de justificar. Por ejemplo, los datos pueden estar en la forma de **rangos**. Esta situación es frecuente en la práctica. Por ejemplo, puede recurrirse a un panel de jueces para evaluar la calidad global de 10 formulaciones diferentes de una bebida refrescante asignando a la "mejor" formulación el rango 1, a la "segunda mejor" formulación el rango 2, y así sucesivamente. Es improbable que los datos ordenados por rangos satisfagan el supuesto de normalidad. Muchos métodos no paramétricos incluyen el análisis de rangos y, por consiguiente, se ajustan idealmente a este tipo de problema.

14-2 PRUEBA DE SIGNOS

14-2.1 Descripción de la prueba

La prueba de signos se usa para probar hipótesis acerca de la mediana $\tilde{\mu}$ de una distribución continua. La mediana de una distribución es el valor de la variable aleatoria X tal que la proba-

bilidad de que un valor observado de X sea menor o igual que la mediana es 0.5, y la probabilidad de que un valor observado de X sea mayor o igual que la mediana es 0.5. Es decir, $P(X \leq \tilde{\mu}) = P(X \geq \tilde{\mu}) = 0.5$.

Puesto que la distribución normal es simétrica, la media de una distribución normal es igual a la mediana. Por lo tanto, la prueba de signos puede usarse para probar hipótesis acerca de la media de una distribución normal. Se trata del mismo problema en el que se utilizó la prueba t en el capítulo 8. En la sección 14-2.4 se revisarán los méritos relativos de los dos procedimientos. Obsérvese que, si bien la prueba t está diseñada para muestras de una distribución normal, la prueba de signos es apropiada para muestras de cualquier distribución continua. Por tanto, la prueba de signos es un procedimiento no paramétrico.

Suponga que las hipótesis son

$$\begin{aligned} H_0 &: \tilde{\mu} = \tilde{\mu}_0 \\ H_1 &: \tilde{\mu} < \tilde{\mu}_0 \end{aligned} \tag{14-1}$$

Es sencillo describir el procedimiento de la prueba. Suponga que $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de la población de interés. Se forman las diferencias

$$X_i - \tilde{\mu}_0, \quad i = 1, 2, \ldots, n \tag{14-2}$$

Ahora bien, si la hipótesis nula H_0: $\tilde{\mu} = \tilde{\mu}_0$ es verdadera, cualquier diferencia $X_i - \tilde{\mu}_0$ tiene la misma posibilidad de ser positiva o negativa. Un estadístico de la prueba apropiado es el número de estas diferencias que son positivas, digamos R^+. Por lo tanto, para probar la hipótesis nula en realidad se prueba que el número de signos positivos es el valor de una variable aleatoria binomial que tiene el parámetro $p = 1/2$. A partir de la distribución binomial puede calcularse directamente un valor P para el número observado de signos positivos r^+. Por ejemplo, al probar las hipótesis de la ecuación 14-1 anterior, se rechazará H_0 en favor de H_1 sólo si la proporción de signos positivos es significativamente menor que 1/2 (o de manera equivalente, siempre que el número observado de signos r^+ sea muy pequeño). Por tanto, si el valor P calculado

$$P = P\left(R^+ \leq r^+ \text{ cuando } p = \frac{1}{2}\right)$$

es menor o igual que algún nivel de significación α preseleccionado, se rechazará H_0 y se concluirá que H_1 es verdadera.

Para probar la otra hipótesis de una cola

$$\begin{aligned} H_0 &: \tilde{\mu} = \tilde{\mu}_0 \\ H_1 &: \tilde{\mu} > \tilde{\mu}_0 \end{aligned} \tag{14-3}$$

se rechazará H_0 en favor de H_1 sólo si el número observado de signos positivos, por ejemplo r^+, es grande o, de manera equivalente, siempre que la proporción observada de signos positivos sea significativamente mayor que 1/2. Por tanto, si el valor P calculado

$$P = P\left(R^+ \geq r^+ \text{ cuando } p = \frac{1}{2}\right)$$

es menor que α, se rechazará H_0 y se concluirá que H_1 es verdadera.

También es posible probar la hipótesis alternativa de dos colas. Si las hipótesis son

$$\begin{aligned} H_0 &: \tilde{\mu} = \tilde{\mu}_0 \\ H_1 &: \tilde{\mu} \neq \tilde{\mu}_0 \end{aligned} \tag{14-4}$$

deberá rechazarse $H_0 : \tilde{\mu} = \tilde{\mu}_0$ si la proporción de signos positivos difiere de manera significativa (porque sea mayor o menor) de 1/2. Esto es equivalente a que el número observado de signos positivos r^+ sea lo suficientemente grande o pequeño. Por tanto, si $r^+ < n/2$, el valor P es

$$P = 2P\left(R^+ \leq r^+ \text{ cuando } p = \frac{1}{2}\right)$$

y si $r^+ > n/2$, el valor P es

$$P = 2P\left(R^+ \geq r^+ \text{ cuando } p = \frac{1}{2}\right)$$

Si el valor P es menor que algún nivel preseleccionado α, se rechazará H_0 y se concluirá que H_1 es verdadera.

EJEMPLO 14-1

Montgomery y Peck reportan un estudio en el que un motor cohete se forma uniendo un propulsor de ignición y un propulsor de sustentación en el interior de una carcasa de metal. La resistencia al corte de la unión entre los dos tipos de propulsores es una característica importante. En la tabla 14-1 se muestran los resultados de las pruebas de 20 motores seleccionados al azar. Quiere probarse la hipótesis de que la mediana del esfuerzo cortante es 2 000 psi, utilizando $\alpha = 0.05$.

Este problema puede resolverse aplicando el procedimiento de prueba de hipótesis de los ocho pasos que se introdujo en el capítulo 8.

1. El parámetro de interés es la mediana de la distribución de la resistencia al corte de los propulsores.
2. $H_0 : \tilde{\mu} = 2\,000$ psi
3. $H_1 : \tilde{\mu} \neq 2\,000$ psi
4. $\alpha = 0.05$
5. El estadístico de la prueba es el número observado de las diferencias positivas de la tabla 14-1, o $r^+ = 14$.

Tabla 14-1 Datos de la resistencia al corte de los propulsores

Observación i	Resistencia al corte x_i	Diferencias $x_i - 2\,000$	Signo
1	2 158.70	+ 158.70	+
2	1 678.15	–321.85	–
3	2 316.00	+ 316.00	+
4	2 061.30	+ 61.30	+
5	2 207.50	+ 207.50	+
6	1 708.30	–291.70	–
7	1 784.70	–215.30	–
8	2 575.10	+ 575.10	+
9	2 357.90	+ 357.90	+
10	2 256.70	+ 256.70	+
11	2 165.20	+ 165.20	+
12	2 399.55	+ 399.55	+
13	1 779.80	–220.20	–
14	2 336.75	+ 336.75	+
15	1 765.30	–234.70	–
16	2 053.50	+ 53.50	+
17	2 414.40	+414.40	+
18	2 200.50	+200.50	+
19	2 654.20	+654.20	+
20	1 753.70	–246.30	–

6. Se rechazará H_0 si el valor P correspondiente a $r^+ = 14$ es menor o igual que $\alpha = 0.05$.
7. Cálculos: Puesto que $r^+ = 14$ es mayor que $n/2 = 20/2 = 10$, el valor P se calcula con

$$\begin{aligned} P &= 2P\left(R^+ \geq 14 \text{ cuando } p = \frac{1}{2}\right) \\ &= 2\sum_{r=14}^{20}\binom{20}{r}(0.5)^r(0.5)^{20-r} \\ &= 0.1153 \end{aligned}$$

8. Conclusiones: puesto que $P = 0.1153$ no es menor que $\alpha = 0.05$, no puede rechazarse la hipótesis nula de que la mediana de la resistencia al corte es 2 000 psi. Otra forma de expresar esto es que el número observado de signos positivos $r^+ = 14$ no fue lo suficientemente grande o pequeño para indicar que la mediana de la resistencia al corte es diferente de 2 000 psi con el nivel de significación $\alpha = 0.05$.

También es posible construir una tabla de valores críticos para la prueba de signos, la cual se muestra en la tabla IX del apéndice. El uso de esta tabla para la hipótesis alternativa de dos colas de la ecuación 14-4 es sencillo. Como anteriormente, sea que R^+ denote el número de diferencias $(X_i - \tilde{\mu}_0)$ que son positivas y sea que R^- denote el número de estas diferencias que son negativas. Sea $R = \text{mín}(R^+, R^-)$. En la tabla IX del apéndice se presentan los valores críticos r_α^* de la prueba de signos que aseguran que P(error tipo I) = P(rechazar H_0 cuando H_0 es verdadera) $= \alpha$ para $\alpha = 0.01$, $\alpha = 0.05$ y $\alpha = 0.10$. Si el valor observado del estadístico de la prueba $r \leq r_\alpha^*$, entonces la hipótesis nula H_0: $\tilde{\mu} = \tilde{\mu}_0$ deberá rechazarse.

Para ilustrar cómo se usa esta tabla, referirse a los datos de la tabla 14-1 que se utilizaron en el ejemplo 14-1. Ahora bien, $r^+ = 14$ y $r^- = 6$; por lo tanto, $r = \text{mín}(14, 6) = 6$. En la tabla IX del apéndice con $n = 20$ y $\alpha = 0.05$, se encuentra que $r_{0.05}^* = 5$. Puesto que $r = 6$ no es menor o igual que el valor crítico $r_{0.05}^* = 5$, no puede rechazarse la hipótesis nula de que la mediana de la resistencia al corte es 2 000 psi.

La tabla IX del apéndice también puede usarse con la prueba de signos cuando la hipótesis alternativa es de una cola. Si la hipótesis alternativa es $H_1 : \tilde{\mu} > \tilde{\mu}_0$, entonces se rechaza $H_0 : \tilde{\mu} = \tilde{\mu}_0$ si $r^- \leq r_\alpha^*$; si la hipótesis alternativa es $H_1 : \tilde{\mu} < \tilde{\mu}_0$, entonces se rechaza $H_0 : \tilde{\mu} = \tilde{\mu}_0$ si $r^+ \leq r_\alpha^*$. El nivel de significación de una prueba de una cola es la mitad del valor de una prueba de dos colas. En la tabla IX del apéndice se muestran los niveles de significación de una cola inmediatamente abajo de los niveles de dos colas en los encabezados de las columnas.

Por último, obsérvese que cuando el estadístico de la prueba tiene una distribución discreta, como R en la prueba de signos, puede ser imposible elegir un valor crítico r_α^* que tenga un nivel de significación exactamente igual a α. El enfoque usado en la tabla IX del apéndice es elegir r_α^* que produzca un valor de α que esté tan próximo como sea posible al nivel de significación α anunciado.

Empates en la prueba de signos

Puesto que se supone que la población fundamental es continua, la probabilidad de encontrar un "empate" es cero; es decir, un valor de X_i exactamente igual a $\tilde{\mu}_0$. Sin embargo, en ocasiones éstos pueden darse en la práctica debido a la forma en que se recaban los datos. Cuando ocurren empates, deberán dejarse de lado y aplicar la prueba de signos a los datos restantes.

La aproximación normal

Cuando $p = 0.5$, una distribución normal es una buena aproximación de la distribución binominal cuando n es al menos 10. Por tanto, puesto que la media de la distribución binomial es np y la varianza es $np(1 - p)$, la distribución de R^+ es aproximadamente normal con media $0.5n$ y varianza $0.25n$ siempre que n sea moderadamente grande. Por lo tanto, en estos casos la hipótesis nula $H_0 : \tilde{\mu} = \tilde{\mu}_0$ puede probarse usando el estadístico

$$Z_0 = \frac{R^+ - 0.5n}{0.5\sqrt{n}} \qquad (14\text{-}5)$$

La hipótesis alternativa de dos colas se rechazaría si el valor observado del estadístico de la prueba $|z_0| > z_{\alpha/2}$, y las regiones críticas de la hipótesis alternativa de una cola se elegirían para reflejar el sentido de la hipótesis alternativa. (Si la hipótesis alternativa es H_1: $\tilde{\mu} > \tilde{\mu}_0$, se rechaza H_0 si $z_0 > z_\alpha$, por ejemplo.)

EJEMPLO 14-2

Se ilustrará el procedimiento de la aproximación normal aplicándolo al problema del ejemplo 14-1. Recuérdese que los datos de este ejemplo se encuentran en la tabla 14-1. El procedimiento de los ocho pasos es el siguiente.

1. El parámetro de interés es la mediana de la distribución de la resistencia al corte de los propulsores.
2. $H_0 : \tilde{\mu} = 2\,000$ psi
3. $H_1 : \tilde{\mu} \neq 2\,000$ psi
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$z_0 = \frac{r^+ - 0.5n}{0.5\sqrt{n}}$$

6. Puesto que $\alpha = 0.05$, se rechazará H_0 en favor de H_1 si $|z_0| > z_{0.025} = 1.96$
7. Cálculos: Puesto que $r^+ = 14$, el estadístico de la prueba es

$$z_0 = \frac{14 - 0.5(20)}{0.5\sqrt{20}} = 1.789$$

8. Conclusiones: puesto que $z_0 = 1.789$ no es mayor que $z_{0.025} = 1.96$, la hipótesis nula no puede rechazarse. Por tanto, las conclusiones coinciden exactamente con las del ejemplo 14-1.

14-2.2 Prueba de signos para muestras pareadas

La prueba de signos también puede aplicarse a observaciones pareadas sacadas de poblaciones continuas. Sea (X_{1j}, X_{2j}), $j = 1, 2, \ldots, n$ una colección de observaciones pareadas de dos poblaciones continuas, y sean

$$D_j = X_{1j} - X_{2j} \quad j = 1, 2, \ldots, n$$

las diferencias pareadas. Quiere probarse la hipótesis de que las dos poblaciones tienen una mediana común, es decir, que $\tilde{\mu}_1 = \tilde{\mu}_2$. Esto es equivalente a probar que la mediana de las diferencias $\tilde{\mu}_D = 0$. Esto puede hacerse aplicando la prueba de signos a las n diferencias observadas d_j, como se ilustra en el siguiente ejemplo.

EJEMPLO 14-3

Un ingeniero automotriz investiga dos tipos diferentes de dispositivos de medición en un sistema electrónico de inyección de combustible a fin de determinar si difieren en el rendimiento de combustible. El sistema se instala en 12 automóviles diferentes, y la prueba se realiza instalando cada dispositivo de medición en cada uno de los automóviles. Los datos observados del rendimiento de combustible, las diferencias correspondientes y sus signos se muestran en la tabla 14-2. Quiere usarse la prueba de signos para determinar si la mediana del rendimiento de combustible es la misma para ambos dispositivos usando $\alpha = 0.05$. Se presenta a continuación el procedimiento de los ocho pasos.

1. Los parámetros de interés son las medianas del rendimiento de combustible de los dos dispositivos de medición.
2. $H_0 : \tilde{\mu}_1 = \tilde{\mu}_2$, o, de manera equivalente, $H_0 : \tilde{\mu}_D = 0$
3. $H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$, o, de manera equivalente, $H_1 : \tilde{\mu}_D \neq 0$
4. $\alpha = 0.05$
5. Se usará la tabla IX del apéndice para realizar la prueba, por lo que el estadístico de la prueba es $r = \text{mín}\,(r^+, r^-)$.
6. Puesto que $\alpha = 0.05$ y $n = 12$, en la tabla IX del apéndice se obtienen los valores críticos $r^*_{0.05} = 2$. Se rechazará H_0 en favor de H_1 si $r \leq 2$.
7. Cálculos: en la tabla 14-2 se presentan las diferencias y sus signos, y se observa que $r^+ = 8$, $r^- = 4$ y, por tanto, $r = \text{mín}(8, 4) = 4$.
8. Conclusiones: puesto que $r = 4$ no es menor o igual que el valor crítico $r^*_{0.05} = 2$, no puede rechazarse la hipótesis nula de que los dos dispositivos producen la misma mediana del rendimiento de combustible.

Tabla 14-2 Desempeño de los dispositivos de medición de flujo

	Dispositivo de medición			
Automóvil	1	2	Diferencia, d_j	Signo
1	17.6	16.8	0.8	+
2	19.4	20.0	–0.6	–
3	19.5	18.2	1.3	+
4	17.1	16.4	0.7	+
5	15.3	16.0	–0.7	–
6	15.9	15.4	0.5	+
7	16.3	16.5	–0.2	–
8	18.4	18.0	0.4	+
9	17.3	16.4	0.9	+
10	19.1	20.1	–1.0	–
11	17.8	16.7	1.1	+
12	18.2	17.9	0.3	+

14-2.3 Error tipo II para la prueba de signos

La prueba de signos controlará la probabilidad del error tipo I en un nivel predeterminado α cuando se prueba la hipótesis nula $H_0 : \tilde{\mu} = \tilde{\mu}_0$ en cualquier distribución continua. Como en cualquier procedimiento de prueba de hipótesis, es importante investigar la probabilidad del error tipo II, β. La prueba deberá tener la capacidad de detectar con efectividad las desviaciones de la hipótesis nula, y una buena medición de esta efectividad es el valor de β para las desviaciones que son importantes. Un valor pequeño de β implica que el procedimiento de prueba es efectivo.

Cuando se determina β, es importante tener presente no sólo que debe usarse un valor particular de $\tilde{\mu}$, por ejemplo $\tilde{\mu}_0 + \Delta$, sino también que la *forma* de la distribución fundamental afectará los cálculos. Para ilustrar, suponer que la distribución fundamental es normal con $\sigma = 1$ y que se está probando la hipótesis $H_0 : \tilde{\mu} = 2$ contra $H_1 : \tilde{\mu} > 2$. (Puesto que $\tilde{\mu} = \mu$ en la distribución normal, esto es equivalente a probar que la media es igual a 2.) Suponer que es importante detectar una desviación de $\tilde{\mu} = 2$ a $\tilde{\mu} = 3$. La situación se ilustra gráficamente en la figura 14-1*a*. Cuando la hipótesis alternativa es verdadera $(H_1 : \tilde{\mu} = 3)$, la probabilidad de que la variable aleatoria X es menor o igual que el valor 2 es

$$p = P(X \le 2) = P(Z \le -1) = \Phi(-1) = 0.1587$$

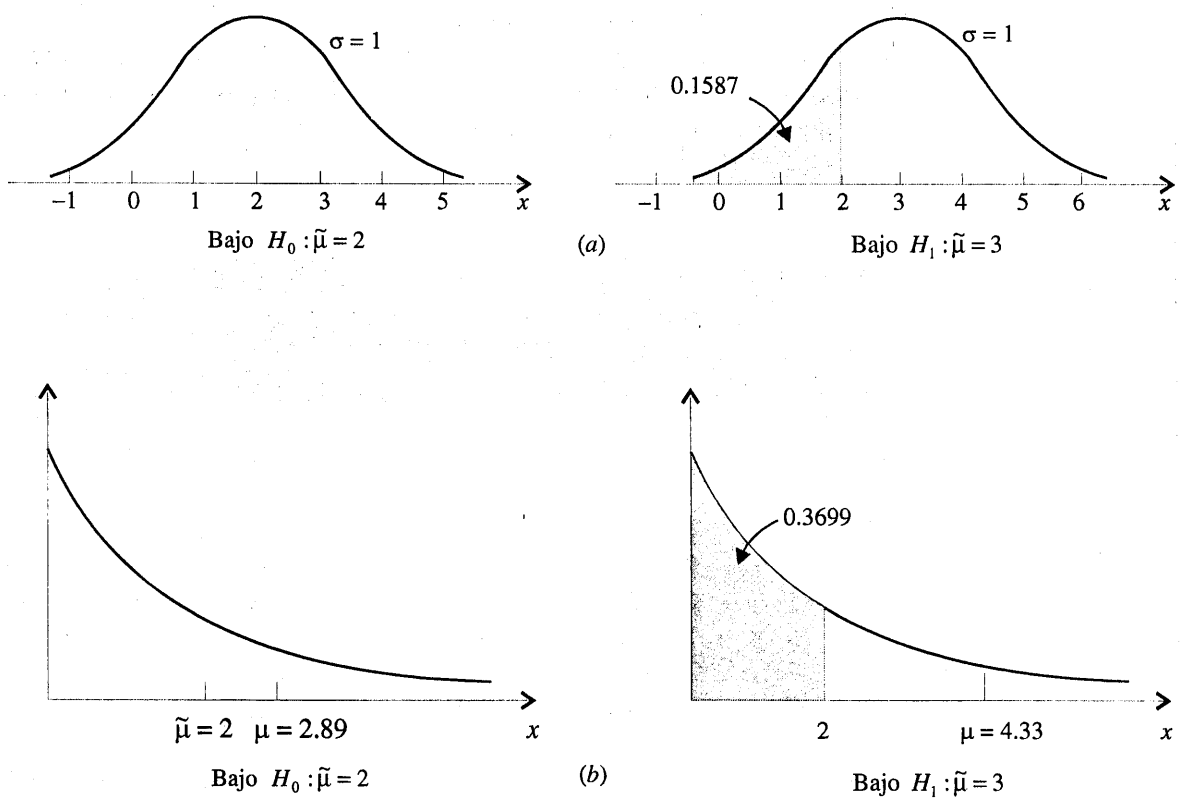


Figura 14-1 Cálculo de β para la prueba de signos. ***a***) Distribuciones normales, ***b***) distribuciones exponenciales.

Suponga que se ha tomado una muestra aleatoria de tamaño 12. En el nivel $\alpha = 0.05$, la tabla IX del apéndice indica que se rechazaría $H_0 : \tilde{\mu} = 2$ si $r^- \leq r^*_{0.05}$. Por lo tanto, β es la probabilidad de que no se rechace $H_0 : \tilde{\mu} = 2$ cuando en realidad $\tilde{\mu} = 3$, o

$$\beta = 1 - \sum_{x=0}^{2} \binom{12}{x} (0.1587)^x (0.8413)^{12-x} = 0.2944$$

Si la distribución de X hubiera sido exponencial en vez de normal, entonces la situación sería como se muestra en la figura 14-1*b*, y la probabilidad de que la variable aleatoria X sea menor o igual que el valor $x = 2$ cuando $\tilde{\mu} = 3$ (obsérvese que cuando la mediana de una distribución exponencial es 3, la media es 4.33) es

$$p = P(X \leq 2) = \int_0^2 \frac{1}{4.33} e^{-\frac{1}{4.33}x} dx = 0.3699$$

En este caso,

$$\beta = 1 - \sum_{x=0}^{2} \binom{12}{x} (0.3699)^x (0.6301)^{12-x} = 0.8794$$

Por tanto, en la prueba de signos β depende no sólo del valor alternativo de $\tilde{\mu}$, sino también del área a la derecha del valor especificado en la hipótesis nula bajo la distribución de probabilidad de la población. Esta área depende en gran medida de la forma de esa distribución de probabilidad particular.

14-2.4 Comparación con la prueba *t*

Si la población fundamental es normal, entonces sería posible ya sea aplicar la prueba de signos o bien la prueba *t* para probar $H_0 : \tilde{\mu} = \tilde{\mu}_0$. La prueba *t* se caracteriza por tener el valor más pequeño de β posible entre todas las pruebas que tienen el nivel de significación α para la hipótesis alternativa de una cola y para las pruebas con regiones críticas simétricas para la hipótesis alternativa de dos colas, por lo que es superior a la prueba de signos en el caso de la distribución normal. Cuando la distribución de la población es simétrica y no normal (pero con media finita $\mu = \tilde{\mu}$), la prueba *t* tendrá una β menor (o una potencia mayor) que la prueba de signos, a menos que la distribución tenga colas muy importantes en comparación con la distribución normal. Por tanto, la prueba de signos suele considerarse más bien como un procedimiento de prueba para la mediana que como un rival serio de la prueba *t*. La prueba de rangos con signo de Wilcoxon que se revisa en la siguiente sección es preferible a la prueba de signos, y es equiparable a la prueba *t* para distribuciones simétricas.

EJERCICIOS DE LA SECCIÓN 14-2

14-1. Se tomaron 10 muestras de un baño de plateado usado en un proceso de manufactura de productos electrónicos, y se determinó el pH del baño. Los valores muestrales del pH son 7.91, 7.85, 6.82, 8.01, 7.46, 6.95, 7.05, 7.35, 7.25, 7.42. El ingeniero de manufactura piensa que el valor de la mediana del pH es 7.0. ¿Los datos muestrales indi-

can que es correcta esta afirmación? Use la prueba de signos con $\alpha = 0.05$ para investigar esta hipótesis. Encuentre el valor P para esta prueba.

14-2. El contenido de titanio en una aleación utilizada en aviones es un factor determinante de la resistencia. Una muestra de 20 elementos para prueba revela el siguiente contenido de titanio (en por ciento):

8.32, 8.05, 8.93, 8.65, 8.25, 8.46, 8.52, 8.35 8.36, 8.41, 8.42, 8.30, 8.71, 8.75, 8.60, 8.83, 8.50, 8.38, 8.29, 8.46

La mediana del contenido de titanio debería ser 8.5%. Use la prueba de signos con $\alpha = 0.05$ para investigar esta hipótesis. Encuentre el valor P para esta prueba.

14-3. El nivel de impurezas (en ppm) de un producto químico intermedio se mide rutinariamente. Los siguientes datos se observaron en una prueba reciente:

2.4, 2.5, 1.7, 1.6, 1.9, 2.6, 1.3, 1.9, 2.0, 2.5, 2.6, 2.3, 2.0, 1.8, 1.3, 1.7, 2.0, 1.9, 2.3, 1.9, 2.4, 1.6

¿Puede afirmarse que la mediana del nivel de impurezas es menor que 2.5 ppm? Enuncie y pruebe la hipótesis apropiada usando la prueba de signos con $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

14-4. Considérense los datos del ejercicio 14-1. Use la aproximación normal de la prueba de signos para probar $H_0 : \tilde{\mu} = 7.0$ contra $H_1 : \tilde{\mu} \neq 7.0$. ¿Cuál es el valor P para esta prueba?

14-5. Considérense los datos de la fuerza de compresión del ejercicio 9-32.

a) Use la prueba de signos para investigar la afirmación de que la mediana de la resistencia a la compresión es de al menos 2 250 psi. Use $\alpha = 0.05$.

b) Use la aproximación normal para probar la misma hipótesis que se formuló en el inciso a). ¿Cuál es el valor P para esta prueba?

14-6. Considérense los datos del contenido de grasa de la margarina del ejercicio 9-31. Use la prueba de signos para probar $H_0 : \tilde{\mu} = 17.0$ contra $H_1 : \tilde{\mu} \neq 17.0$ usando $\alpha = 0.05$. Encuentre el valor P para el estadístico de la prueba y use esta cantidad para llegar a una conclusión.

14-7. Considérense los datos del ejercicio 14-2. Use la aproximación normal de la prueba de signos para probar $H_0 : \tilde{\mu} = 8.5$ contra $H_1 : \tilde{\mu} \neq 8.5$ usando $\alpha = 0.05$. ¿Cuál es el valor P para esta prueba?

14-8. Considérense los datos del ejercicio 14-3. Use la aproximación normal de la prueba de signos para probar $H_0 : \tilde{\mu} = 2.5$ contra $H_1 : \tilde{\mu} < 2.5$. ¿Cuál es el valor P para esta prueba?

14-9. En un probador de dureza Rockwell pueden usarse dos tipos diferentes de puntas. Se seleccionaron 8 muestras de metal de prueba de una aleación basada en níquel, y cada muestra se probó dos veces, una con cada punta. Las lecturas de la dureza en la escala C de Rockwell se muestran en la siguiente tabla. Use la prueba de signos con $\alpha = 0.05$ para determinar si las dos puntas producen o no lecturas de la dureza equivalentes.

Muestra de metal	Punta 1	Punta 2
1	63	60
2	52	51
3	58	56
4	60	59
5	55	58
6	57	54
7	53	52
8	59	61

14-10. Dos formulaciones diferentes de una base para pintura pueden usarse en acondicionadora de aluminio. El tiempo de secado de estas dos formulaciones es una consideración importante en el proceso de manufactura. Se seleccionan 20 entrepaños; la mitad de cada entrepaño se pinta con la base uno y la otra mitad se pinta con la base dos. Se registran los tiempos de secado, los cuales se reportan en la siguiente tabla. ¿Hay evidencia de que la mediana de los tiempos de

secado de las dos formulaciones sean diferentes? Use la prueba de signos con $\alpha = 0.01$.

	Tiempo de secado (en hr)	
Panel	Formulación 1	Formulación 2
1	1.6	1.8
2	1.3	1.5
3	1.5	1.5
4	1.6	1.7
5	1.7	1.6
6	1.9	2.0
7	1.8	2.1
8	1.6	1.7
9	1.4	1.6
10	1.8	1.9
11	1.9	2.0
12	1.8	1.9
13	1.7	1.5
14	1.5	1.7
15	1.6	1.6
16	1.4	1.2
17	1.3	1.6
18	1.6	1.8
19	1.5	1.6
20	1.8	2.0

14-11. Use la aproximación normal de la prueba de signos para los datos del ejercicio 14-10. ¿Qué conclusiones pueden sacarse?

14-12. El diámetro de un rodamiento de bolas fue medido por 12 inspectores, para lo cual cada uno de ellos utilizó dos tipos diferentes de calibradores. Los resultados fueron los siguientes:

Inspector	Calibrador 1	Calibrador 2
1	0.265	0.264
2	0.265	0.265
3	0.266	0.264
4	0.267	0.266
5	0.267	0.267
6	0.265	0.268
7	0.267	0.264
8	0.267	0.265
9	0.265	0.265
10	0.268	0.267
11	0.268	0.268
12	0.265	0.269

¿Hay alguna diferencia significativa entre las medianas de la población de las mediciones representada por las dos muestras? Use $\alpha = 0.05$.

14-13. Considérense los datos del colesterol en la sangre del ejercicio 9-35. Use la prueba de signos para determinar si hay alguna diferencia entre las medianas de los dos grupos de mediciones usando $\alpha = 0.05$. ¿Qué conclusión práctica se sacaría de este estudio?

14-14. Use la aproximación normal de la prueba de signos para los datos del ejercicio 14-12. Con $\alpha = 0.05$, ¿qué conclusiones pueden sacarse?

14-15. Use la aproximación normal de la prueba de signos para los datos del ejercicio 14-13. Con $\alpha = 0.05$, ¿qué conclusiones pueden sacarse?

14-16. La distribución del tiempo entre los accesos a un sistema de telecomunicaciones es exponencial, y el administrador del sistema quiere probar la hipótesis de que H_0: $\tilde{\mu} = 3.5$ min contra $H_1 : \tilde{\mu} > 3.5$ min.

a) ¿Cuál es el valor de la media de la distribución exponencial bajo $H_0 : \tilde{\mu} = 3.5$?

b) Suponga que se ha tomado una muestra *de* $n = 10$ observaciones y que se obtiene $r^- = 3$. ¿Se rechazaría H_0 en la prueba de signos con $\alpha = 0.05$?

c) ¿Cuál es la probabilidad del error tipo II en esta prueba si $\tilde{\mu} = 4.5$?

14-17. Suponga que se toma una muestra de $n = 10$ mediciones de una distribución normal con $\sigma = 1$. Quiere probarse $H_0 : \mu = 0$ contra $H_0 : \mu = 0$. El estadístico normal de la prueba es $Z_0 = \bar{X} / \left(\sigma / \sqrt{n}\right)$, y se decide usar una región crítica de 1.96 (es decir, se rechaza H_0 si $z_0 \geq 1.96$).

a) ¿Qué valor tiene α en esta prueba?

b) ¿Qué valor tiene β en esta prueba, si $\mu = 1$?

c) Si se usa la prueba de signos, especifique la región crítica que da un valor de α consistente con α de la prueba normal.

d) ¿Cuál es el valor de β en la prueba de signos, si $\mu = 1$? Compare este valor con el resultado obtenido en el inciso b).

14-18. Considérese el estadístico de la prueba en la prueba de signos del ejercicio 14-9. Encuentre el valor P para este estadístico.

14-19. Considérese el estadístico de la prueba en la prueba de signos del ejercicio 14-10. Encuentre el valor P para este estadístico. Compárelo con el valor P para el estadístico de la prueba de la aproximación normal calculado en el ejercicio 14-11.

14-3 PRUEBA DE RANGOS CON SIGNO DE WILCOXON

La prueba de signos utiliza tan sólo los signos positivos y negativos de las diferencias entre las observaciones y la mediana $\tilde{\mu}_0$ (o los signos positivos y negativos de las diferencias entre las observaciones en el caso de muestras pareadas). No toma en consideración el tamaño o magnitud de estas diferencias. Frank Wilcoxon ideó un procedimiento de prueba que utiliza tanto la dirección (el signo) como la magnitud. Este procedimiento, conocido hoy como la **prueba de rangos con signo de Wilcoxon**, se analiza y ejemplifica en esta sección.

La prueba de rangos con signo de Wilcoxon se aplica al caso de **distribuciones continuas y simétricas**. Bajo estos supuestos, la media es igual a la mediana, y puede usarse este procedimiento para probar la hipótesis nula de que $\mu = \mu_0$. Se indica a continuación cómo hacerlo.

14-3.1 Descripción de la prueba

El interés se encuentra en probar H_0: $\mu = \mu_0$ contra las hipótesis alternativas usuales. Suponga que $X_1, X_2, ..., X_n$ es una muestra aleatoria de una distribución continua y simétrica con media (y mediana) μ. Se calculan las diferencias $X_i - \mu_0$, $i = 1, 2, ..., n$. Se ordenan las diferencias absolutas $|X_i - \mu_0|$, $i = 1, 2, ..., n$, en orden ascendente, y después a los rangos o posiciones ordenadas se les asignan los signos de sus diferencias correspondientes. Sea W^+ la suma de rangos positivos y sea W^- el valor absoluto de la suma de rangos negativos, y sea $W = \text{mín}(W^+, W^-)$. La tabla X del apéndice contiene los valores críticos de W, por ejemplo w_α^*. Si la hipótesis alternativa es $H_1 : \mu \neq \mu_0$, entonces si el valor observado del estadístico $w \leq w_\alpha^*$, se rechaza la hipótesis nula $H_0 : \mu = \mu_0$. La tabla X del apéndice incluye los niveles de significación de $\alpha = 0.10$, $\alpha = 0.05$, $\alpha = 0.02$, $\alpha = 0.01$ para la prueba de dos colas.

Para las pruebas de una cola, si la hipótesis alternativa es $H_1 : \mu > \mu_0$, se rechaza $H_0 : \mu = \mu_0$ si $w^- \leq w_\alpha^*$; y si la hipótesis alternativa es $H_1 : \mu < \mu_0$, se rechaza $H_0 : \mu = \mu_0$ si $w^+ \leq w_\alpha^*$. Los niveles de significación para las pruebas de una cola que se incluyen en la tabla X del apéndice son $\alpha = 0.05$, 0.025, 0.01 y 0.005.

EJEMPLO 14-4

La prueba de rangos con signo de Wilcoxon se ilustrará aplicándola a los datos de la resistencia al corte de los propulsores de la tabla 14-1. Suponga que la distribución fundamental es continua y simétrica. El procedimiento de los ocho pasos se aplica de la manera siguiente:

1. El parámetro de interés es la media (o la mediana) de la distribución de la resistencia al corte de los propulsores.
2. $H_0 : \mu = 2\,000$ psi

3. $H_1 : \mu \neq 2\,000$ psi
4. $\alpha = 0.05$
5. El estadístico de prueba es

$$w = \text{mín}\ (w^+, w^-)$$

6. Se rechazará H_0 si $w \leq w^*_{0.05} = 52$, por la tabla X del apéndice.
7. Cálculos: los rangos con signo de la tabla 14-1 se muestran en la siguiente tabla.

Observación	Diferencia $X_i - 2\,000$	Rango con signo
16	+53.50	+1
4	+61.30	+2
1	+158.70	+3
11	+165.20	+4
18	+200.50	+5
5	+207.50	+6
7	−215.30	−7
13	−220.20	−8
15	−234.70	−9
20	−246.30	−10
10	+256.70	+11
6	−291.70	−12
3	+316.00	+13
2	−321.85	−14
14	+336.75	+15
9	+357.90	+16
12	+399.55	+17
17	+414.40	+18
8	+575.10	+19
19	+654.20	+20

La suma de rangos positivos es $w^+ = (1 + 2 + 3 + 4 + 5 + 6 + 11 + 13 + 15 + 16 + 17 + 18 + 19 + 20) = 150$, y la suma de los valores absolutos de los rangos negativos es $w^- = (7 + 8 + 9 + 10 + 12 + 14) = 60$. Por lo tanto

$$w = \text{mín}\ (150, 60) = 60$$

8. Conclusiones: puesto que $w = 60$ no es menor o igual que el valor crítico $w_{0.05} = 52$, no puede rechazarse la hipótesis nula de que la media (o la mediana, pues se ha supuesto que la población es simétrica) de la resistencia al corte 2 000 psi.

Empates en la prueba de rangos con signo de Wilcoxon

Como la población fundamental continua, en teoría los empates son imposibles, aun cuando en ocasiones ocurrirán en la práctica. Si varias observaciones tienen la misma magnitud absoluta, se asignan al promedio de los rangos o posiciones que tomarían si difirieran ligeramente entre sí.

14-3.2 Aproximación para muestras grandes

Si el tamaño de la muestra es moderadamente grande, por ejemplo $n > 20$, entonces puede demostrarse que W^+ (o W^-) tiene aproximadamente una distribución normal con media

$$\mu_{W+} = \frac{n(n+1)}{4}$$

y varianza

$$\sigma^2_{W+} = \frac{n(n+1)(2n+1)}{24}$$

Por lo tanto, la prueba de $H_0 : \mu = \mu_0$ puede basarse en el estadístico

$$Z_0 = \frac{W^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \tag{14-6}$$

En una tabla de la distribución normal estándar puede escogerse una región crítica apropiada para las hipótesis alternativas ya sea de una o bien de dos colas.

14-3.3 Observaciones pareadas

La prueba de rangos con signo de Wilcoxon puede aplicarse a datos pareados. Sea (X_{1j}, X_{2j}), $j = 1, 2, ..., n$, una colección de observaciones pareadas de dos distribuciones continuas que difieren únicamente con respecto a sus medias. (No es necesario que las distribuciones de X_1 y X_2 sean simétricas.) Con esto se asegura que la distribución de las diferencias $D_j = X_{1j} - X_{2j}$ es continua y simétrica. Por tanto, la hipótesis nula es $H_0 : \mu_1 = \mu_2$, que es equivalente a $H_0 : \mu_D = 0$. Se considera inicialmente la hipótesis alternativa de dos colas $H_1 : \mu_1 \neq \mu_2$ (o $H_1 : \mu_D \neq 0$).

Para usar la prueba de rangos con signo de Wilcoxon, primero se ordenan las diferencias de sus valores absolutos en orden ascendente, y después a los rangos se les asignan los signos de las

diferencias. Los empates se asignan a rangos promedio. Sea W^+ la suma de rangos positivos y W^- el valor absoluto de la suma de rangos negativos, y sea $W = \text{mín}(W^+, W^-)$. Si el valor observado $w \leq w_\alpha^*$, la hipótesis nula $H_0 : \mu_1 = \mu_2$ (o $H_0 : \mu_D = 0$) se rechaza, donde w_α^* se elige en la tabla X del apéndice.

Para pruebas de una cola, si la hipótesis alternativa es $H_1 : \mu_1 > \mu_2$ (o $H_1 : \mu_D > 0$), se rechaza H_0 si $w^- \leq w_\alpha^*$; y si $H_1 : \mu_1 < \mu_2$ (o $H_1 : \mu_D < 0$), se rechaza H_0 si $w^+ \leq w_\alpha^*$. Asegúrese de usar los niveles de significación de la prueba de una cola mostrados en la tabla X del apéndice.

EJEMPLO 14-5

Se aplicará la prueba de rangos con signo de Wilcoxon a los datos de la prueba del dispositivo para medir el combustible usados previamente en el ejemplo 14-3. El procedimiento de prueba de hipótesis de los ocho pasos puede aplicarse de la siguiente manera:

1. Los parámetros de interés son las medias del rendimiento de combustible de los dos dispositivos de medición.
2. $H_0 : \mu_1 = \mu_2$ o, de manera equivalente, $H_0 : \mu_D = 0$
3. $H_1 : \mu_1 \neq \mu_2$ o, de manera equivalente, $H_1 : \mu_D \neq 0$
4. $\alpha = 0.05$
5. El estadístico de la prueba es

$$w = \text{mín}\ (w^+, w^-)$$

   donde w^+ y w^- son las sumas de los rangos positivos y negativos de las diferencias de la tabla 14-2.
6. Puesto que $\alpha = 0.05$ y $n = 12$, en la tabla X del apéndice se obtiene el valor crítico $w_{0.05}^* = 13$. Se rechazará $H_0 : \mu_D = 0$ si $w \leq 13$.
7. Cálculos: Utilizando los datos de la tabla 14-2, se calculan los siguientes rangos con signo.

Automóvil	Diferencia	Rango con signo
7	−0.2	−1
12	0.3	2
8	0.4	3
6	0.5	4
2	−0.6	−5
4	0.7	6.5
5	−0.7	−6.5
1	0.8	8
9	0.9	9
10	−1.0	−10
11	1.1	11
3	1.3	12

Obsérvese que $w^+ = 55.5$ y $w^- = 22.5$. Por lo tanto

$$w = \text{mín}\,(55.5, 22.5) = 22.5$$

8. Conclusiones: Puesto que $w = 22.5$ no es menor o igual que $w^*_{0.05} = 13$, no puede rechazarse la hipótesis nula de que los dos dispositivos de medición producen el mismo rendimiento de combustible.

14-3.4 Comparación con la prueba *t*

Cuando la población fundamental es normal, puede usarse la prueba *t* o la prueba de rangos con signo de Wilcoxon para probar hipótesis sobre μ. Como ya se mencionó, la prueba *t* es la mejor prueba en tales situaciones en el sentido de que produce un valor mínimo de β para todas las pruebas con nivel de significación α. Sin embargo, puesto que no siempre se tiene la seguridad que la distribución normal es apropiada, y como en muchas situaciones ésta es inapropiada, existe interés en comparar los dos procedimientos para poblaciones tanto normales como no normales.

Desafortunadamente, esta comparación no es sencilla. El problema radica en que es muy complicado obtener el valor de β para la prueba de rangos con signo de Wilcoxon, y es difícil obtener β para la prueba *t* en el caso de distribuciones no normales. Debido a las dificultades para comparar los errores tipo II, se han desarrollado otras medidas de comparación. Una medida de uso generalizado es la **eficiencia asintótica relativa** (EAR). La EAR de una prueba en comparación con otra es el cociente límite de los tamaños de las muestras que se necesitan para obtener probabilidades de error idénticas en los dos procedimientos. Por ejemplo, si la EAR de una prueba en comparación relativa con una prueba alternativa es 0.5, entonces cuando los tamaños de las muestras son grandes, la primera prueba requerirá una muestra del doble del tamaño de la segunda para obtener una representación del error similar. Aun cuando esto no hace referencia alguna a tamaños de las muestras pequeños, puede decirse lo siguiente:

1. Para poblaciones normales, la EAR de la prueba de rangos con signo de Wilcoxon en comparación con la prueba *t* es aproximadamente 0.95.
2. Para poblaciones no normales, la EAR es al menos 0.86, y en muchos casos excederá la unidad.

Aun cuando estos son resultados para muestras grandes, se concluye en general que la prueba de rangos con signo de Wilcoxon nunca será mucho peor que la prueba *t* y que en muchos casos en que la población no es normal puede ser superior. Por tanto, la prueba de rangos con signo de Wilcoxon es una alternativa útil para la prueba *t*.

EJERCICIOS DE LA SECCIÓN 14-3

14-20. Considérense los datos del ejercicio 14-1 y suponer que la distribución del pH es simétrica y continua. Use la prueba de rangos con signo de Wilcoxon con $\alpha = 0.05$ para probar la hipótesis $H_0 : \mu = 7$ contra $H_1 : \mu \neq 7$.

14-21. Considérense los datos del ejercicio 14-2. Supnga que la distribución del conteni-

do de titanio es simétrica y continua. Use la prueba de rangos con signo de Wilcoxon con $\alpha = 0.05$ para probar la hipótesis $H_0 : \mu = 8.5$ contra $H_1 : \mu \neq 8.5$.

14-22. Considérense los datos del ejercicio 14-2. Use la aproximación para muestras grandes de la prueba del rango con signo de Wilcoxon para probar la hipótesis H_0: $\mu = 8.5$ contra $H_1 : \mu \neq 8.5$. Use $\alpha = 0.05$. Suponga que la distribución del contenido de titanio es continua y simétrica.

14-23. Considérense los datos del ejercicio 14-3. Use la prueba de rangos con signo de Wilcoxon para probar la hipótesis $H_0 : \mu = 2.5$ ppm contra $H_1 : \mu < 2.5$ ppm usando $\alpha = 0.05$. Suponga que la distribución del nivel de impurezas es continua y simétrica.

14-24. Considérense los datos de la prueba de dureza Rockwell del ejercicio 14-9. Suponga que ambas distribuciones son continuas y use la prueba de rangos con signo de Wilcoxon para probar que la diferencia media en las lecturas de la dureza entre las dos puntas es cero. Use $\alpha = 0.05$.

14-25. Considérense los datos del tiempo de secado de la pintura del ejercicio 14-10. Suponga que ambas poblaciones son continuas, y use la prueba de rangos con signo de Wilcoxon para probar que la diferencia en los tiempos medios de secado entre las dos formulaciones es cero. Use $\alpha = 0.01$.

14-26. Aplique la prueba de rangos con signo de Wilcoxon a los datos de las mediciones del ejercicio 14-12. Use $\alpha = 0.05$ y suponga que las dos distribuciones de las mediciones son continuas.

14-27. Aplique la prueba de rangos con signo de Wilcoxon a los datos del colesterol en la sangre a los que se hace referencia en el ejercicio 14-13. Use $\alpha = 0.05$ y suponga que las dos distribuciones son continuas.

14-4 PRUEBA DE LA SUMA DE RANGOS DE WILCOXON

Suponga que se tienen dos poblaciones continuas e independientes X_1 y X_2 con medias μ_1 y μ_2. Suponga que las distribuciones de X_1 y X_2 tienen la misma forma y dispersión, y que difieren tan sólo (posiblemente) en sus localizaciones. La prueba de la suma de rangos de Wilcoxon puede usarse para probar la hipótesis H_0: $\mu_1 = \mu_2$. A este procedimiento se le llama en ocasiones la prueba Mann-Whitney, aun cuando el estadístico de la prueba Mann-Whitney por lo general se expresa en una forma diferente.

14-4.1 Descripción de la prueba

Sean $X_{11}, X_{12}, \ldots, X_{1n_1}$ y $X_{21}, X_{22}, \ldots, X_{2n_2}$ dos muestras aleatorias independientes de tamaños $n_1 \leq n_2$ de las poblaciones continuas X_1 y X_2 descritas arriba. Quieren probarse las hipótesis

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

El procedimiento de prueba es el siguiente. Se ordenan todas las $n_1 + n_2$ observaciones en orden de magnitud ascendente y se les asignan rangos. Si dos o más observaciones están empatadas (son idénticas), entonces se usa la media de los rangos que se les hubieran asignado si las

observaciones difirieran. Sea W_1 la suma de rangos de la muestra menor (1), y se define W_2 como la suma de rangos de la otra muestra. Entonces,

$$W_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - W_1 \tag{14-7}$$

Ahora bien, si las medias muestrales no difieren, se esperará que la suma de rangos sea aproximadamente igual para ambas muestras después de hacer el ajuste para la diferencia en el tamaño de las muestras. Por consiguiente, si hay una diferencia considerable en las sumas de los rangos, se concluirá que las medias no son iguales.

La tabla XI del apéndice contiene el valor crítico de las sumas de los rangos para $\alpha = 0.05$ y $\alpha = 0.01$, suponiendo la hipótesis alternativa de dos colas anterior. Al referirse a la tabla XI del apéndice con los tamaños de las muestras apropiados n_1 y n_2, podrá obtenerse el valor crítico w_α. La hipótesis nula $H_0 : \mu_1 = \mu_2$ se rechaza en favor de $H_1 : \mu_1 \neq \mu_2$ si cualquiera de los valores observados w_1 o w_2 es menor o igual que el valor crítico w_α de la tabla.

El procedimiento también puede usarse para hipótesis alternativas de una cola. Si la hipótesis alternativa es $H_1 : \mu_1 < \mu_2$, entonces se rechaza H_0 si $w_1 \leq w_\alpha$; para $H_1 : \mu_1 > \mu_2$, se rechaza H_0 si $w_2 \leq w_\alpha$. Para estas pruebas de una cola, los valores críticos w_α de la tabla corresponden a los niveles de significación de $\alpha = 0.025$ y $\alpha = 0.005$.

EJEMPLO 14-6

Se estudia el esfuerzo axial promedio en los miembros de tensión usados en una estructura aeronáutica. Se encuentran bajo estudio dos aleaciones. La aleación 1 es un material tradicional y la aleación 2 es una nueva aleación de aluminio y litio que es mucho más ligera que el material estándar. Se prueban 10 ejemplares de cada tipo de aleación, y se mide el esfuerzo axial. Los datos muestrales se presentan en la tabla 14-3. Utilizando $\alpha = 0.05$, quiere probarse la hipótesis de que las medias de las dos distribuciones del esfuerzo son idénticas.

Se aplicará el procedimiento de prueba de hipótesis de los ocho pasos en este problema.

1. Los parámetros de interés son las medias de las dos distribuciones del esfuerzo axial.
2. $H_0 : \mu_1 = \mu_2$
3. $H_1 : \mu_1 \neq \mu_2$
4. $\alpha = 0.05$
5. Se Usará el estadístico de la prueba de la suma de rangos de Wilcoxon de la ecuación 14-7,

$$w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - w_1$$

Tabla 14-3 Esfuerzo axial de dos aleaciones de aluminio y litio

Aleación 1		Aleación 2	
238 psi	3 254 psi	3 261 psi	3 248 psi
3 195	3 229	3 187	3 215
3 246	3 225	3 209	3 226
3 190	3 217	3 212	3 240
3 204	3 241	3 258	3 234

6. Puesto que $\alpha = 0.05$ y $n_1 = n_2 = 10$, en la tabla XI del apéndice se obtiene el valor crítico $w_{0.05} = 78$. Si w_1 o w_2 es menor o igual que $w_{0.05} = 78$, se rechazará $H_0 : \mu_1 = \mu_2$.
7. Cálculos: los datos de la tabla 14-3 se analizan en orden ascendente y se asignan posiciones de la siguiente manera:

Número de aleación	Esfuerzo axial	Rango
2	3 187 psi	1
1	3 190	2
1	3 195	3
1	3 204	4
2	3 209	5
2	3 212	6
2	3 215	7
1	3 217	8
1	3 225	9
2	3 226	10
1	3 229	11
2	3 234	12
1	3 238	13
2	3 240	14
1	3 241	15
1	3 246	16
2	3 248	17
1	3 254	18
2	3 258	19
2	3 261	20

La suma de rangos de la aleación 1 es

$$w_1 = 2+3+4+8+9+11+13+15+16+18 = 99$$

y para la aleación 2

$$w_2 = \frac{(n_1+n_2)(n_1+n_2+1)}{2} - w_1 = \frac{(10+10)(10+10+1)}{2} - 99 = 111$$

8. Conclusiones: puesto que ni w_1 ni w_2 es menor o igual que $w_{0.05} = 78$, no puede rechazarse la hipótesis nula de que ambas aleaciones presentan el mismo esfuerzo axial promedio.

14-4.2 Aproximación con muestras grandes

Cuando tanto n_1 como n_2 son moderadamente grandes, por ejemplo mayores que 8, una aproximación adecuada de la distribución de w_1 puede ser la distribución normal con media

$$\mu_{W_1} = \frac{n_1(n_1+n_2+1)}{2}$$

y varianza

$$\sigma^2_{W_1} = \frac{n_1 n_2 (n_1+n_2+1)}{12}$$

Por lo tanto, para n_1 y $n_2 > 8$, podría usarse

$$Z_0 = \frac{W_1 - \mu_{W_1}}{\sigma_{W_1}} \tag{14-8}$$

como estadístico, y la región crítica apropiada es $|z_0| > z_{\alpha/2}$, $z_0 > z_\alpha$, o $z_0 < -z_\alpha$, dependiendo de si la prueba es de dos colas, de cola superior o de cola inferior.

14-4.3 Comparación con la prueba *t*

En la sección 14-3.4 se presentó la comparación de la prueba *t* con la prueba de rangos con signo de Wilcoxon. Los resultados para el problema de dos muestras son idénticos al caso de una muestra. Es decir, cuando el supuesto de normalidad es correcto, la prueba de la suma de rangos de Wilcoxon tiene aproximadamente 95% de la eficiencia de la prueba *t* para muestras grandes. Por otra parte, independientemente de la forma de las distribuciones, la prueba de la suma de rangos de Wilcoxon tendrá siempre una eficiencia de al menos 86%. La eficiencia de la prueba de Wilcoxon en comparación con la prueba *t* suele ser alta si la distribución fundamental tiene colas más importantes que la distribución normal, debido a que el comportamiento de la prueba *t* depende en gran medida de la media muestral, la cual es bastante inestable en distribuciones con colas importantes.

EJERCICIOS DE LA SECCIÓN 14-4

14-28. Un ingeniero eléctrico debe diseñar un circuito que entregue la cantidad máxima de corriente a un cinescopio para conseguir una imagen con la brillantez suficiente. Dentro de sus restricciones de diseño permisibles, el ingeniero ha desarrollado dos circuitos candidatos y ha probado prototipos de cada uno. Los datos resultantes (en microamperes) son los siguientes:

Circuito 1:	251, 255, 258, 257, 250, 251, 254, 250, 248
Circuito 2:	250, 253, 249, 256, 259, 252, 260 251

Use la prueba de la suma de rangos de Wilcoxon para probar $H_0: \mu_1 = \mu_2$ contra la hipótesis alternativa $H_1: \mu_1 > \mu_2$. Use $\alpha = 0.025$.

14-29. El autor realiza viajes regulares a Seattle, Washington, en su calidad de consultor de

Boeing Company. Utiliza una de dos líneas aéreas, Delta y Alaska. En ocasiones las demoras de los vuelos son inevitables, pero el autor estaría dispuesto a contratar la mayor parte de sus vuelos con la línea aérea con el mejor registro de llegadas a tiempo. El número de minutos que sus vuelos llegaron retrasados en los seis últimos viajes en cada línea aérea se muestran a continuación. ¿Hay evidencia de que alguna de las dos líneas aéreas tenga un desempeño superior en cuanto a las llegadas a tiempo? Use $\alpha = 0.01$ y la prueba de la suma de rangos de Wilcoxon.

Delta:	13, 10, 1, – 4, 0, 9 (minutos de retraso)
Alaska:	15, 8, 3, – 1, –2, 4 (minutos de retraso)

14-30. Al fabricante de tubos incandescentes le interesa probar dos elementos de calentamiento diferentes para su producto. El elemento que produzca la ganancia máxima de calor después de 15 minutos sería preferible. Obtiene 10 muestras de cada unidad de calentamiento y prueba cada una de ellas. La ganancia de calor después de 15 minutos (en °F) se muestra a continuación. ¿Hay alguna razón para sospechar que una unidad es superior a la otra? Use $\alpha = 0.05$ y la prueba de la suma de rangos de Wilcoxon.

Unidad 1	25, 27, 29, 31, 30, 26, 24, 32, 33, 38
Unidad 2	31, 33, 32, 35, 34, 29, 38, 35, 37, 30

14-31. Use la aproximación normal de la prueba de la suma de rangos de Wilcoxon en el problema del ejercicio 14-28. Suponga que $\alpha = 0.05$. Encuentre el valor P aproximado para este estadístico de la prueba.

14-32. Use la aproximación normal de la prueba de la suma de rangos de Wilcoxon en el experimento de la ganancia de calor del ejercicio 14-30. Suponga que $\alpha = 0.05$. ¿Cuál es el valor P aproximado para este estadístico de la prueba?

14-33. Considérense los datos de la rapidez de grabado químico del ejercicio 9-19. Use la prueba de la suma de rangos de Wilcoxon para investigar la afirmación de que la rapidez de grabado promedio es la misma para ambas soluciones. Si $\alpha = 0.05$, ¿a qué conclusiones se llega?

14-34. Use la prueba de la suma de rangos de Wilcoxon en el experimento de la temperatura de deflexión de los tubos que se describe en el ejercicio 9-18. Si $\alpha = 0.05$, ¿a qué conclusiones se llega?

14-35. Use la aproximación normal de la prueba de la suma de rangos de Wilcoxon en el problema del ejercicio 9-19. Suponga que $\alpha = 0.05$. Encuentre el valor P aproximado para esta prueba.

14-36. Use la aproximación normal de la prueba de la suma de rangos de Wilcoxon en el problema del ejercicio 9-18. Suponga que $\alpha = 0.05$. Encuentre el valor P aproximado para esta prueba.

14-5 MÉTODOS NO PARAMÉTRICOS EN EL ANÁLISIS DE VARIANZA

14-5.1 Prueba de Kruskal-Wallis

El modelo para el análisis de varianza con un solo factor que se desarrolló en el capítulo 12 para comparar a medias poblacionales es

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n_i \end{cases} \tag{14-9}$$

En este modelo, se supone que los términos del error ε_{ij} tienen una distribución normal e independiente con media cero y varianza σ_2. El supuesto de normalidad llevó directamente a la prueba F descrita en el capítulo 12. La prueba de Kruskal-Wallis es una alternativa no paramétrica de la prueba F; se requiere únicamente que las ε_{ij} tengan la misma distribución continua para todos los niveles del factor $i = 1, 2, ..., a$.

Suponga que $N = \sum_{i=1}^{a} n_i$ es el número total de observaciones. Se ordenan las N observaciones de la menor a la mayor, y a la observación menor se le asigna el rango 1, a la siguiente más pequeña el rango 2, ..., y a la observación más grande el rango N. Si la hipótesis nula

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$

es verdadera, las N observaciones provienen de la misma distribución, y todas las asignaciones posibles de los N rangos a las a muestras son igualmente posibles, entonces podría esperarse que los rangos o posiciones 1, 2, ..., N estuvieran mezclados entre las a muestras. Sin embargo, si la hipótesis nula H_0 es falsa, entonces algunas muestras constarán de observaciones que tienen predominantemente rangos pequeños, mientras que otras muestras constarán de observaciones que tienen rangos predominantemente grandes. Sea R_{ij} el rango de la observación Y_{ij}, y sea que $R_{i\cdot}$ y $\bar{R}_{i\cdot}$ denoten el total y el promedio de los n_i rangos del tratamiento i-ésimo. Cuando la hipótesis nula es verdadera, entonces

$$E(R_{ij}) = \frac{N+1}{2}$$

y

$$E(\bar{R}_{i\cdot}) = \frac{1}{n_i}\sum_{j=1}^{n_i} E(R_{ij}) = \frac{N+1}{2}$$

El estadístico de la prueba de Kruskal-Wallis mide el grado en que los rangos promedio $\bar{R}_{i\cdot}$ observados en realidad, difieren de su valor esperado $(N + 1)/2$. Si esta diferencia es grande, entonces la hipótesis nula H_0 se rechaza. El estadístico de la prueba es

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{a} n_i\left(\bar{R}_{i\cdot} - \frac{N+1}{2}\right)^2 \qquad (14\text{-}10)$$

Una fórmula de cálculo alternativa que en ocasiones resulta más conveniente es

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{a} \frac{R_{i\cdot}^2}{n_i} - 3(N+1) \qquad (14\text{-}11)$$

En general se preferiría la ecuación 14-11 a la ecuación 14-10, porque incluye los totales de los rangos en lugar de los promedios.

La hipótesis nula H_0 deberá rechazarse si los datos muestrales generan un valor grande de H. La distribución nula de H se ha obtenido usando el hecho de que bajo H_0 cada asignación posible de los rangos a los a tratamientos es igualmente posible. Por tanto, podrían enumerarse todas las asignaciones posibles y contar el número de veces que ocurre cada valor de H. Esto ha llevado a tablas de los valores críticos de H, aun cuando la mayoría de las tablas están restringidas a tamaños de muestras pequeños n_i. En la práctica, por lo general se emplea la siguiente aproximación para muestras grandes: Siempre que H_0 es verdadera y se cumple alguna de las condiciones

$$a = 3 \quad \text{y} \quad n_i \geq 6 \quad \text{para } i = 1, 2, 3$$
$$a > 3 \quad \text{y} \quad n_i \geq 5 \quad \text{para } i = 1, 2, \ldots, a$$

entonces H tiene aproximadamente una distribución ji-cuadrada con $a - 1$ grados de libertad. Puesto que valores grandes de H implican que H_0 es falsa, se rechazará H_0 si el valor observado

$$h \geq \chi^2_{\alpha, a-1}$$

La prueba tiene un nivel de significación aproximado α.

Empates en la prueba de Kruskal-Wallis

Cuando hay observaciones empatadas, se asigna un rango promedio a cada una de las observaciones empatadas. Cuando hay empates, el estadístico de la prueba de la ecuación 14-11 deberá reemplazarse por

$$H = \frac{1}{S^2}\left[\sum_{i=1}^{a} \frac{R_{i\cdot}^2}{n_i} - \frac{N(N+1)^2}{4}\right] \tag{14-12}$$

donde n_i es el número de observaciones del tratamiento i-ésimo, N es el número total de observaciones y

$$S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{a} \sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4}\right] \tag{14-13}$$

Obsérvese que S^2 no es sino la varianza de los rangos. Cuando el número de empates es moderado, habrá poca diferencia entre las ecuaciones 14-11 y 14-12 y puede usarse la forma más simple (la ecuación 14-11).

EJEMPLO 14-7

Montgomery presentó datos de un experimento en el que se probaron cinco niveles diferentes del contenido de algodón en una fibra sintética para determinar si el contenido de algodón tiene algún efecto sobre la resistencia a la tensión de la fibra. Los datos muestrales y los rangos de este experimento se muestran en la tabla 14-4. Se aplicará la prueba de Kruskal-Wallis a estos datos usando $\alpha = 0.01$.

Tabla 14-4 Datos y rangos del experimento para probar la resistencia a la tensión

Porcentaje de algodón							$r_{i\cdot}$
15	y_{1j}	7	7	15	11	9	
rangos	r_{1j}	2.0	2.0	12.5	7.0	4.0	27.5
20	y_{2j}	12	17	12	18	18	
rangos	r_{2j}	9.5	14.0	9.5	16.5	16.5	66.0
25	y_{3j}	14	18	18	19	19	
rangos	r_{3j}	11.0	16.5	16.5	20.5	20.5	85.0
30	y_{4j}	19	25	22	19	23	
rangos	r_{4j}	20.5	25.0	23.0	20.5	24.0	113.0
35	y_{5j}	7	10	11	15	11	
rangos	r_{5j}	2.0	5.0	7.0	12.5	7.0	33.5

Puesto que hay un número bastante grande de empates, se usa la ecuación 14-12 como estadístico de la prueba. Por la ecuación 14-13 se encuentra

$$S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} r_{ij}^2 - \frac{N(N+1)^2}{4}\right]$$
$$= \frac{1}{24}\left[5510 - \frac{25(26)^2}{2}\right]$$
$$= 53.54$$

y el estadístico de la prueba es

$$h = \frac{1}{s^2}\left[\sum_{i=1}^{a}\frac{r_{i\cdot}^2}{n_i} - \frac{N(N+1)^2}{4}\right]$$
$$= \frac{1}{53.54}\left[5\,245.7 - \frac{25(26)^2}{4}\right]$$
$$= 19.06$$

Puesto que $h > \chi^2_{0.01,4} = 13.28$, se rechazaría la hipótesis nula y se concluiría que los tratamientos difieren. Se llega a la misma conclusión con el análisis de varianza de la prueba F usual.

14-5.2 Transformación de rangos

Al procedimiento usado en la sección anterior mediante el cual las observaciones son reemplazadas por sus rangos se le llama la **transformación de rangos**. Es una técnica muy poderosa y

de uso generalizado. Si se aplicara la prueba F común a los rangos en vez de a los datos originales, se obtendría

$$F_0 = \frac{H/(a-1)}{(N-1-H)(N-a)}$$

como el estadístico de la prueba. Obsérvese que, cuando el estadístico H de la prueba de Kruskal-Wallis se incrementa o se decrementa, también F_0 se incrementa o se decrementa. Ahora bien, puesto que la distribución F_0 es aproximada por la distribución F, la prueba de Kruskal-Wallis es aproximadamente equivalente al aplicar el análisis de varianza usual a los rangos.

La transformación de rangos tiene diversas aplicaciones en los problemas de diseño experimental en los que no se cuenta con alguna alternativa no paramétrica para el análisis de varianza. Si los datos están ordenados en rangos y se aplica la prueba F ordinaria, resulta un procedimiento aproximado, pero que tiene buenas propiedades estadísticas. Cuando haya incertidumbre en cuanto al supuesto de normalidad o al efecto de puntos atípicos o valores "extraños", se recomienda que el análisis de varianza usual se realice tanto en los datos originales como en los rangos. Cuando ambos procedimientos producen resultados similares, los supuestos del análisis de varianza probablemente se satisfacen razonablemente bien, y el análisis estándar es satisfactorio. Cuando los dos procedimientos difieren, deberá preferirse la transformación de rangos, ya que es menos posible que sea distorsionado por la falta de normalidad y las observaciones inusuales. En tales casos, el experimentador quizá quiera investigar el uso de transformaciones para condiciones de no normalidad y examinar los datos y el procedimiento experimental para determinar si están presentes puntos atípicos y por qué han ocurrido.

EJERCICIOS DE LA SECCIÓN 14-5

14-37. Montgomery presentó los resultados de un experimento para comparar cuatro técnicas de mezclado diferentes sobre la resistencia de tensión del cemento portland. Los resultados se muestran en la siguiente tabla. ¿Hay algún indicio de que la técnica de mezclado afecta la resistencia? Use $\alpha = 0.05$.

Técnica de mezclado	Resistencia a la tensión (lb/pulg^2)			
1	3 129	3 000	2 865	2 890
2	3 200	3 000	2 975	3 150
3	2 800	2 900	2 985	3 050
4	2 600	2 700	2 600	2 765

14-38. En un artículo de *Quality Control Handbook*, 3a. edición (McGraw-Hill) se presentan los resultados de un experimento realizado para investigar el efecto de tres diferentes métodos de acondicionamiento sobre la resistencia a la ruptura de aglomerados (briquetas) de cemento. Los datos se muestran en la siguiente tabla. Usando $\alpha = 0.05$, ¿hay algún indicio de que el método de acondicionamiento afecta la resistencia a la ruptura?

Método de acondicionamiento	Resistencia a la ruptura (lb/pulg^2)				
1	553	550	568	541	537
2	553	599	579	545	540
3	492	530	528	510	571

14-39. En *Statistics for Research* (John Wiley & Sons), S. Dowdy y S. Wearden presentaron los resultados de un experimento para medir el desempeño de sierras de mano. En el experimento se midió el ángulo de contragolpe en que se desvía la sierra cuando empieza a cortar un tablón sintético de 3 pul-

gadas. En la siguiente tabla se muestran los ángulos de desviación de cinco sierras elegidas al azar de cada uno de cuatro fabricantes. ¿Hay evidencia de que los productos de los fabricantes difieren con respecto al ángulo de contragolpe? Use $\alpha = 0.01$.

Fabricante	Ángulo de contragolpe				
A	42	17	24	39	43
B	28	50	44	32	61
C	57	45	48	41	54
D	29	40	22	34	30

14-40. Considérense los datos del ejercicio 12-2. Use el procedimiento de Kruskal-Wallis con $\alpha = 0.05$ para probar las diferencias entre la uniformidad promedio con las tres velocidades del flujo del gas.

14-41. Encuentre el valor *P* aproximado del estadístico de la prueba calculado en el ejercicio 14-37.

14-42. Encuentre el valor *P* aproximado del estadístico de la prueba calculado en el ejercicio 14-40.

Ejercicios complementarios

14-43. El terminado superficial de 10 partes metálicas producidas en un proceso de esmerilado es el siguiente (en μpulg): 10.32, 9.68, 9.92, 10.10, 10.20, 9.87, 10.14, 9.74, 9.80, 10.26. ¿Los datos apoyan la afirmación de que el valor de la mediana del terminado superficial es 10 μpulg? Use la prueba de signos con $\alpha = 0.05$. ¿Cuál es el valor *P* para esta prueba?

14-44. Use la aproximación normal de la prueba de signos en el problema del ejercicio 14-43. Encuentre el valor *P* para esta prueba. ¿Cuáles son las conclusiones si $\alpha = 0.05$?

14-45. Las emisiones de fluoruro (en ppm) de una planta química se monitorean rutinariamente. Las siguientes son 15 observaciones basadas en muestras de aire tomadas al azar durante un mes de producción: 7, 3, 4, 2, 5, 6, 9, 8, 7, 3, 4, 4, 3, 2, 6. ¿Puede afirmarse que la mediana del nivel de impurezas de fluoruro es menor que 6 ppm? Enuncie y pruebe las hipótesis apropiadas usando la prueba de signos con $\alpha = 0.05$. ¿Cuál es el valor *P* de esta prueba?

14-46. Use la aproximación normal de la prueba de signos en el problema del ejercicio 14-45. ¿Cuál es el valor *P* para esta prueba?

14-47. Considérense los datos del ejercicio 9-38. Use la prueba de signos con $\alpha = 0.05$ para determinar si hay alguna diferencia en la mediana de las lecturas de las impurezas entre las dos pruebas analíticas.

14-48. Considérense los datos del ejercicio 14-43. Use la prueba de rangos con signo de Wilcoxon en este problema con $\alpha = 0.05$. ¿Cuáles son las hipótesis que se están probando en este problema?

14-49. Considérense los datos del ejercicio 14-45. Use la prueba de rangos con signo de Wilcoxon en este problema con $\alpha = 0.05$. ¿Qué conclusiones pueden sacarse? ¿La hipótesis que se está probando aquí difiere de la hipótesis que se probó originalmente en el ejercicio 14-45?

14-50. Use la prueba de rangos con signo de Wilcoxon con $\alpha = 0.05$ en el experimento de la modificación de la dieta descrito en el ejercicio 9-37. Enuncie cuidadosamente las conclusiones que pueden sacarse de este experimento.

14-51. Use la prueba de la suma de rangos con signo de Wilcoxon $\alpha = 0.01$ para el estudio de la economía de combustible descrito en el ejercicio 9-75. ¿Qué conclusiones pueden sacarse acerca de la diferencia en la media del rendimiento en millas por galón de los dos vehículos de este estudio?

14-52. Use la aproximación para muestras grandes de la prueba de la suma de rangos de Wilcoxon para los datos de la economía de combustible del ejercicio 9-75. ¿Qué conclusiones pueden sacarse acerca de la diferencia en las medias si $\alpha = 0.01$? Encuentre el valor *P* para esta prueba.

14-53. Use la prueba de la suma de rangos de Wilcoxon con $\alpha = 0.025$ para el experimento de la capacidad de llenado descrito en el ejercicio 9-78. ¿Qué conclusiones pueden sacarse acerca de la capacidad de los dos equipos de llenado?

14-54. Use la aproximación para muestras grandes de la prueba de la suma de rangos de Wilcoxon con $\alpha = 0.025$ para el experimento de la capacidad de llenado descrito en el ejercicio 9-78. Encuentre el valor P para esta prueba. ¿Qué conclusiones pueden sacarse?

14-55. Considérese el experimento de la resistencia de contacto del ejercicio 12-46. Use la prueba de Kruskal-Wallis para probar las diferencias en la resistencia de contacto medio entre las tres aleaciones. Si $\alpha = 0.01$, ¿cuáles son las conclusiones? Encuentre el valor P para esta prueba.

14-56. Considérese el experimento descrito en el ejercicio 12-49. Use la prueba de Kruskal-Wallis en este experimento con $\alpha = 0.05$. ¿Qué conclusiones se sacarían? Encuentre el valor P para esta prueba.

14-57. Considérese el experimento de la calidad del pan del ejercicio 12-50. Use la prueba de Kruskal-Wallis con $\alpha = 0.01$ para analizar los datos de este experimento. Encuentre el valor P para esta prueba. ¿Qué conclusiones pueden sacarse?

EJERCICIOS PARA DESARROLLAR EL INTELECTO

14-58. En la aproximación para muestras grandes de la prueba de rangos con signo de Wilcoxon, deduzca la media y la desviación estándar del estadístico de la prueba usado en el procedimiento.

14-59. **Prueba de tendencias**. El rodete de un turbocargador se fabrica utilizando un proceso de fusión del revestimiento. La flecha se inserta en la abertura del rodete, y esta abertura del rodete es una dimensión crítica. Cuando se forman los moldes de cera, la herramienta que los hace se desgasta. Esto puede ocasionar el agrandamiento de la dimensión de la abertura del rodete. Diez mediciones de la abertura del rodete, hechas en el orden en el tiempo de la producción, son (en mm): 4.00, 4.02, 4.03, 4.01, 4.00, 4.03, 4.04, 4.02, 4.03, 4.03.

a) Suponga que p es la probabilidad de que la observación X_{i+5} exceda la observación X_i. Si no hay ninguna tendencia hacia arriba o hacia abajo, entonces X_{i+5} no es ni más ni menos posible que exceda X_i o que se localice abajo de X_i. ¿Cuál es el valor de p?

b) Sea V el número de valores de i para los que $X_{i+5} > X_i$. Si no hay ninguna tendencia hacia arriba o hacia abajo en las mediciones, ¿cuál es la distribución de probabilidad de V?

c) Utilice los datos anteriores y los resultados de los incisos a) y (b) para probar H_0: no hay ninguna tendencia contra H_1: hay una tendencia hacia arriba. Use $\alpha = 0.05$. Obsérvese que esta prueba es una modificación de la prueba de signos. Fue desarrollada por Cox y Stuart.

14-60. Considérese la prueba de rangos con signo de Wilcoxon y suponga que $n = 5$. Suponga que H_0: $\mu = \mu_0$ es verdadera.

a) ¿Cuántas secuencias diferentes de rangos con signo son posibles? Enumere estas secuencias.

b) ¿Cuántos valores diferentes de W^+ hay? Encuentre la probabilidad asociada con cada valor de W^+.

c) Suponga que la región crítica de la prueba para rechazar H_0 se define como $w^+ > w^*_\alpha$ y $w^*_\alpha = 13$. ¿Cuál es el nivel α aproximado de esta prueba?

d) ¿Este ejercicio indica cómo se desarrollaron los valores críticos para la prueba de rangos con signo de Wilcoxon? Explique su respuesta.

15 Control estadístico de la calidad

PLAN GENERAL DEL CAPÍTULO

15-1 LA ESTADÍSTICA Y EL MEJORAMIENTO DE LA CALIDAD

15-2 CONTROL ESTADÍSTICO DE LA CALIDAD

15-3 CONTROL ESTADÍSTICO DE PROCESOS

15-4 INTRODUCCIÓN A LAS CARTAS DE CONTROL

- 15-4.1 Principios básicos
- 15-4.2 Diseño de una carta de control
- 15-4.3 Subgrupos racionales
- 15-4.4 Análisis de patrones en cartas de control

15-5 CARTAS DE CONTROL $\overline{X}$ Y R

15-6 CARTAS DE CONTROL PARA MEDICIONES INDIVIDUALES

15-7 CAPACIDAD DE UN PROCESO

15-8 CARTAS DE CONTROL DE ATRIBUTOS

- 15-8.1 Carta P (carta de control de proporciones)
- 15-8.2 Carta U (carta de control de defectos por unidad)

15-9 DESEMPEÑO DE LAS CARTAS DE CONTROL

15-10 CARTA DE CONTROL DE SUMA ACUMULADA

15-11 OTRAS HERRAMIENTAS DEL CONTROL ESTADÍSTICO DE PROCESOS PARA RESOLVER PROBLEMAS

15-12 IMPLEMENTACIÓN DEL CONTROL ESTADÍSTICO DE PROCESOS

15-1 LA ESTADÍSTICA Y EL MEJORAMIENTO DE LA CALIDAD

La calidad de los productos y los servicios se ha convertido en un factor de decisión importante en la mayoría de los negocios del mundo actual. Independientemente de si el consumidor es un individuo, una corporación, un programa de defensa militar o una tienda minorista, cuando el consumidor está haciendo una decisión de compra, es posible que asigne igual importancia a la calidad que al costo y al tiempo de entrega. Por consiguiente, el **mejoramiento de la calidad** se ha convertido en una preocupación principal para muchas corporaciones. Este capítulo trata del **control estadístico de la calidad**, un conjunto de herramientas que son esenciales en las actividades de mejoramiento de calidad.

Calidad significa **buena condición para el uso**. Por ejemplo, uno puede adquirir un automóvil con la expectativa de que no tenga defectos de fabricación y que proporcionará un transporte confiable y económico; un minorista compra bienes terminados con la expectativa de que estén empacados y preparados apropiadamente para almacenarlos y exhibirlos con facilidad, o un fabricante compra materia prima y espera procesarla sin problemas de reprocesamiento o de unidades defectuosas. En otras palabras, todos los consumidores esperan que los productos y servicios que adquieren satisfarán sus requerimientos. Esos requerimientos definen la buena condición para el uso.

La calidad o buena condición para el uso se determina mediante la interacción de la **calidad de diseño** y la **calidad de conformidad**. Por calidad de diseño se entienden los diferentes grados o niveles de desempeño, confiabilidad, duración y funcionamiento que son el resultado de decisiones de ingeniería y administrativas deliberadas. Por calidad de conformidad se entiende la **reducción de la variabilidad** y la **eliminación de defectos** sistemáticas, hasta que cada unidad producida sea idéntica y sin defectos.

En nuestra sociedad existe cierta confusión acerca de lo que es el mejoramiento de la calidad; algunas personas siguen pensando que significa darle a un producto una presentación de lujo o gastar más dinero para desarrollar un producto o proceso. Esta concepción está errada. Mejoramiento de calidad significa **eliminar el desperdicio** sistemáticamente. Ejemplos de desperdicio incluyen las piezas de desecho y las tareas de reprocesamiento en las manufacturas, la inspección y las pruebas, los errores en documentos (tales como dibujos de ingeniería, cheques, órdenes de compra y planos), las líneas telefónicas directas para quejas de clientes, el costo de las garantías, y el tiempo requerido para hacer de nuevo cosas que pudieron haberse hecho bien la primera vez. Un esfuerzo de mejoramiento de calidad exitoso puede eliminar gran parte de este desperdicio y redundar en costos menores, una productividad más alta, clientes más satisfechos, la consolidación de la reputación del negocio, una mayor participación en el mercado y, en última instancia, en mayores ganancias para la compañía.

Los métodos estadísticos desempeñan un papel fundamental en el mejoramiento de calidad. A continuación se citan algunas aplicaciones:

1. En el diseño y desarrollo de productos, los métodos estadísticos, incluyendo los experimentos diseñados, pueden usarse para comparar diferentes materiales, componentes o ingredientes, y ayudar a determinar las tolerancias tanto del sistema como de los componentes. Esta aplicación puede bajar de manera significativa los costos de desarrollo y reducir el tiempo de desarrollo.

2. Los métodos estadísticos pueden usarse para determinar la capacidad de un proceso de manufactura. El control estadístico de procesos puede usarse para mejorar de manera sistemática un proceso al reducir la variabilidad.
3. Los métodos de diseño experimental pueden usarse para investigar las mejoras en los procesos. Estas mejoras pueden dar lugar a rendimientos más altos y a costos de manufactura más bajos.
4. Las pruebas de duración proporcionan confiabilidad y datos adicionales sobre el desempeño del producto. Esto puede llevar a diseños nuevos y mejorados, así como a productos con vidas útiles más largas y costos de operación y mantenimiento más bajos.

Algunas de estas aplicaciones se han ilustrado en capítulos anteriores de este libro. Es esencial que ingenieros, científicos y administradores tengan una comprensión profunda de estas herramientas estadísticas en cualquier industria o negocio que quiera ser un productor de alta calidad y bajo costo. En este capítulo se presenta una introducción a los métodos básicos del control estadístico de calidad que, junto con el diseño experimental, constituyen la base de un esfuerzo de mejoramiento de calidad exitoso.

15-2 CONTROL ESTADÍSTICO DE LA CALIDAD

El campo del control estadístico de la calidad puede definirse en un sentido amplio como aquellos métodos estadísticos y de ingeniería que se usan para medir, monitorear, controlar y mejorar la calidad. El control estadístico de la calidad es un campo relativamente nuevo, el cual se remonta a los años veinte. El Dr. Walter A. Shewhart de los laboratorios Bell Telephone fue uno de los pioneros del campo. En 1924 escribió un memorándum que incluía una carta de control moderna, una de las herramientas básicas del control estadístico de procesos. Harold F. Dodge y Harry G. Romig, otros dos empleados de Bell System, aportaron gran parte del liderazgo en el desarrollo de métodos de muestreo e inspección basados en la estadística. El trabajo de estos tres hombres forma gran parte de la base del campo moderno del control estadístico de la calidad. En la Segunda Guerra Mundial se introdujeron de manera generalizada estos métodos en la industria estadounidense. A partir de esa época el Dr. W. Edwards Deming y el Dr. Joseph M. Juran han contribuido a la difusión de los métodos de control estadístico de la calidad.

Los japoneses han sido particularmente exitosos en el despliegue de los métodos de control estadístico de la calidad y han usado métodos estadísticos para conseguir una ventaja significativa sobre sus competidores. En los años setenta, la industria estadounidense sufrió extensivamente la competencia japonesa (y de otras naciones); esto ha llevado, a su vez, al interés renovado en los métodos de control estadístico de la calidad en Estados Unidos. Gran parte de este interés se enfoca en el *control estadístico de procesos* y el *diseño experimental.* Muchas compañías estadounidenses han iniciado amplios programas para implementar estos métodos en sus organizaciones manufactureras, de ingeniería y de otras áreas de negocios.

15-3 CONTROL ESTADÍSTICO DE PROCESOS

No es práctico inspeccionar la calidad dentro de un producto: el producto debe hacerse correctamente la primera vez. En consecuencia, el proceso de fabricación debe ser estable o repetible

y tener la capacidad de operar con poca variabilidad en torno de la dimensión objetivo o nominal. El control estadístico de procesos en línea constituye una poderosa herramienta para conseguir la estabilidad de los procesos y para mejorar su capacidad mediante la reducción de la variabilidad.

Es común considerar el **control estadístico de procesos (CEP)** como un conjunto de herramientas para resolver problemas que puede aplicarse a cualquier proceso. Las principales herramientas del CEP[1] son

1. El histograma
2. El diagrama de Pareto
3. El diagrama de causa y efecto
4. El diagrama de concentración de defectos
5. La carta de control
6. El diagrama de dispersión
7. La hoja de verificación

Aun cuando estas herramientas son una parte importante del CEP, abarcan únicamente el aspecto técnico del tema. Un elemento igualmente importante del CEP es la actitud: el deseo de todos los integrantes de una organización para mejorar en forma continua la calidad y la productividad mediante la reducción sistemática de la variabilidad. La carta de control es la más poderosa de las herramientas del CEP.

15-4 INTRODUCCIÓN A LAS CARTAS DE CONTROL

15-4.1 Principios básicos

En cualquier proceso de producción, sin importar lo bien diseñado que esté o la atención que se preste a su mantenimiento, siempre existirá cierta variabilidad natural o inherente. Esta variabilidad natural o "ruido de fondo" es el efecto acumulado de pequeñas causas en esencia inevitables. Cuando el ruido de fondo de un proceso es relativamente reducido, por lo general se considera que se encuentra en un nivel aceptable del desempeño del proceso. En el contexto del control estadístico de la calidad, a esta variabilidad natural suele llamársele "sistema estable de causas fortuitas o aleatorias". Se dice que un proceso está bajo control estadístico cuando opera únicamente en presencia de **causas fortuitas** de variación. En otras palabras, las causas fortuitas son una parte inherente del proceso.

Ocasionalmente otros tipos de variabilidad pueden estar presentes en la producción de un proceso. Esta variabilidad en las características clave de la calidad por lo general surge de tres fuentes: máquinas mal calibradas, errores del operador o materias primas defectuosas. En general, esta variabilidad es grande cuando se compara con el ruido de fondo, y suele representar un nivel inaceptable del desempeño del proceso. Se hace referencia a estas fuentes de variabilidad

[1] Algunos autores prefieren incluir los métodos de diseño experimental revisados anteriormente como parte del juego de herramientas del CEP. Nosotros no lo hicimos así, porque consideramos el CEP como un enfoque de mejoramiento de la calidad en línea que utiliza técnicas basadas en la observación pasiva del proceso, mientras que el diseño de experimentos es un enfoque activo en el que se hacen cambios deliberados en las variables del proceso. Por ello, con frecuencia se hace referencia a los experimentos diseñados como control de calidad fuera de línea.

que no son parte del patrón de las causas fortuitas como **causas asignables**. Se dice que un proceso que está operando en presencia de causas asignables está fuera de control.[2]

Con frecuencia los procesos de producción operarán en el estado bajo control, produciendo productos aceptables durante periodos relativamente largos. Sin embargo, ocasionalmente se presentarán causas asignables, aparentemente al azar, las cuales dan lugar a un "corrimiento", a un estado fuera de control en el que una gran proporción de la producción del proceso no está conforme con los requerimientos. Uno de los objetivos principales del control estadístico de procesos es detectar con rapidez la ocurrencia de causas asignables o corrimientos del proceso, para que pueda hacerse la investigación del proceso y aplicarse las acciones correctivas antes de que se fabriquen muchas unidades no conformables. La carta de control es una técnica de monitoreo en línea de uso común para este fin.

Recuérdese lo siguiente del capítulo 1. Las figuras 1-17*a* y *b* ilustran que los ajustes de las causas comunes de la variación incrementan la variación de un proceso mientras que las figuras 1-18*a* y *b* ilustran que las acciones deberán tomarse en respuesta a las causas asignables de la variación. Las cartas de control también pueden usarse para estimar los parámetros de un proceso de producción y, gracias a esta información, para determinar la capacidad de un proceso para cumplir con las especificaciones. La carta de control también puede proporcionar información que es útil para mejorar el proceso. Por último, recuérdese que la meta final del control estadístico de procesos es la *eliminación de la variabilidad del proceso*. Aun cuando quizá no es posible eliminar por completo la variabilidad, la carta de control ayuda a reducirla en la medida de lo posible.

En la figura 15-1 se muestra una carta de control típica, que es una representación gráfica de una característica de la calidad que se ha medido o calculado de una muestra contra el número de muestra o el tiempo. Frecuentemente, las muestras se seleccionan en intervalos periódicos, como por ejemplo cada hora. La carta contiene una línea central (LC) que representa el valor promedio de la característica de la calidad correspondiente al estado bajo control. (Es decir, sólo están presentes causas fortuitas.) Otras dos líneas horizontales, llamados el límite de control superior (LCS) y el límite de control inferior (LCI), también se muestran en la carta. Estos

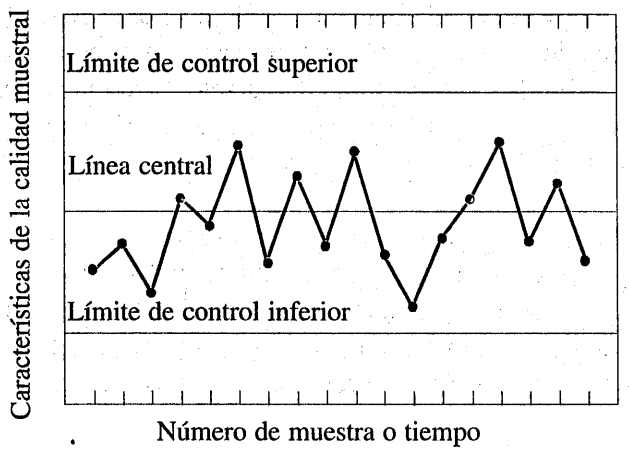


Figura 15-1 Carta de control típica.

[2] La terminología *causas fortuitas* y *causas asignables* fue desarrollada por el Dr. Walter Shewhart. Actualmente, algunos autores usan causa *común* en lugar de causa *fortuita* y causa *especial* en lugar de causa *asignable*.

límites de control se escogen de tal modo que si el proceso está bajo control, casi todos los puntos muestrales se localizarán entre ellos. En general, en tanto los puntos se localicen dentro de los límites de control, se supone que el proceso está bajo control, y que no es necesaria ninguna acción. Sin embargo, un punto que se localice fuera de los límites de control se interpreta como evidencia de que el proceso está fuera de control, y que se requiere investigación y acciones correctivas para encontrar y eliminar la causa o causas asignables responsables de este comportamiento. Los puntos muestrales de la carta de control suelen unirse con segmentos de recta, para que sea más sencillo visualizar cómo ha evolucionado la secuencia de puntos con el tiempo.

Incluso cuando todos los puntos se localicen dentro de los límites de control, si tienen un comportamiento sistemático o no aleatorio, se trata entonces de un indicio de que el proceso está fuera de control. Por ejemplo, si 18 de los últimos 20 puntos se localizan arriba de la línea central pero abajo del límite de control superior y sólo dos de estos puntos se localizan abajo de la línea central pero arriba del límite de control inferior, se tendrían serias sospechas de que algo salió mal. Si el proceso está bajo control, todos los puntos graficados deberían seguir un patrón esencialmente aleatorio. En las cartas de control pueden aplicarse métodos diseñados para encontrar secuencias o patrones no aleatorios como una ayuda para detectar condiciones fuera de control. Un patrón no aleatorio particular aparece con frecuencia en una carta de control por alguna razón, y si esta razón puede encontrarse y eliminarse, el desempeño del proceso podrá mejorarse.

Hay una estrecha conexión entre las cartas de control y la prueba de hipótesis. En esencia, la carta de control es una prueba de la hipótesis de que el proceso se encuentra en un estado de control estadístico. La graficación de un punto dentro de los límites de control es equivalente a no poder rechazar la hipótesis del control estadístico, y un punto que se localice fuera de los límites de control es equivalente a rechazar la hipótesis del control estadístico.

Puede darse un *modelo* general para una carta de control. Sea W un estadístico muestral que mide alguna característica de la calidad de interés, y suponer que la media de W es μ_W y que la desviación estándar de W es σ_W.[3] Entonces la línea central, el límite de control superior y el límite de control inferior son

$$\begin{aligned} UCL &= \mu_W + k\sigma_W \\ CL &= \mu_W \\ LCL &= \mu_W - k\sigma_W \end{aligned} \tag{15-1}$$

donde k es la "distancia" de la línea central a los límites de control, expresada en unidades de desviación estándar. Una elección común es $k = 3$. Esta teoría general de las cartas de control fue propuesta por primera vez por el Dr. Walter A. Shewhart, y las cartas de control que se desarrollan de acuerdo con estos principios suelen llamarse **cartas de control de Shewhart**.

[3] Obsérvese que "sigma" se refiere a la desviación estándar del estadístico graficado en la carta (i.e., σ_W), no a la desviación estándar de la característica de la calidad.

La carta de control es un recurso para describir exactamente qué se entiende por control estadístico; como tal, puede usarse en una variedad de formas. En muchas aplicaciones, se utiliza el monitoreo en línea de procesos. Es decir, se recaban datos muestrales y se emplean para construir la carta de control, y si los valores muestrales de $\bar{x}$ (por ejemplo) quedan dentro de los límites de control y no presentan ningún patrón sistemático, se dice que el proceso está bajo control en el nivel indicado por la carta. Obsérvese que aquí el interés puede ser determinar *tanto* si los datos pasados provinieron de un proceso que estaba bajo control *como* si las muestras futuras de este proceso indican algún control estadístico.

El uso más importante de una carta de control es para *mejorar* el proceso. Se ha encontrado que, en general:

1. La mayoría de los procesos no operan en un estado de control estadístico.
2. Por consiguiente, el uso regular y cuidadoso de cartas de control identificarán las causas asignables. Si estas causas pueden eliminarse del proceso, la variabilidad se reducirá y el proceso mejorará.

Esta actividad de mejoramiento de procesos utilizando la carta de control se ilustra en la figura 15-2. Obsérvese que:

3. La carta de control tan sólo *detectará* las causas asignables. Por lo general será necesaria la *acción* de la administración, del operador y del departamento de ingeniería para eliminar la causa asignable. Es vital un plan de acción para responder a las señales de la carta de control.

Al identificar y eliminar las causas asignables, es importante encontrar la **causa de raíz** del problema y atacarla. Una solución cosmética no redundará en un mejoramiento real del proceso a largo plazo. El desarrollo de un sistema eficaz de acciones correctivas es un componente esencial de una implementación eficaz del CEP.

La carta de control también puede usarse como un *recurso de estimación*. Es decir, a partir de una carta de control que indica control estadístico pueden estimarse ciertos parámetros del proceso, tales como la media, la desviación estándar y la proporción de unidades no conformables o incidentales. Estas estimaciones pueden usarse después para determinar la *capacidad* del pro-

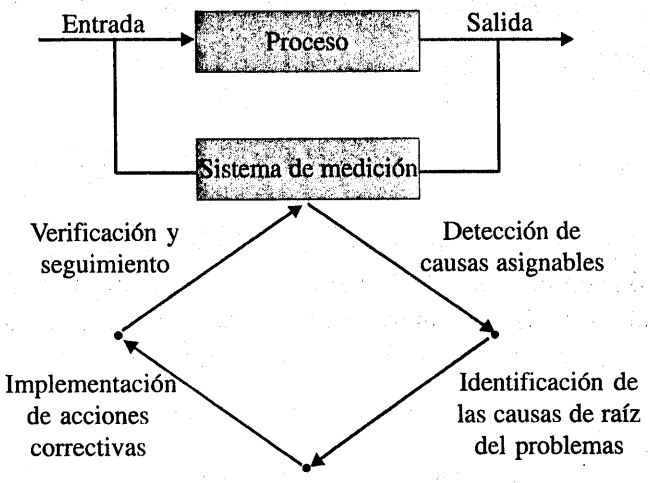


Figura 15-2 Mejoramiento de un proceso utilizando la canta de control.

ceso para producir productos aceptables. Estos **estudios de la capacidad de un proceso** tienen un impacto considerable sobre muchos de los problemas de las decisiones administrativas que ocurren en el ciclo del producto, incluyendo las decisiones sobre fabricar o comprar un componente, las mejoras en la planta y los procesos que reduzcan la variabilidad, y los acuerdos contractuales con clientes o proveedores respecto de la calidad del producto.

Las cartas de control pueden clasificarse en dos tipos generales. Muchas características de la calidad pueden medirse y expresarse como números en una escala continua de medición. En tales casos, es conveniente describir la característica de la calidad con una medida de tendencia central y una medida de variabilidad. Las cartas de control de tendencia central y variabilidad se llaman colectivamente **cartas de control de variables**. La carta $\overline{X}$ es la de uso más generalizado para monitorear la tendencia central, mientras que las cartas basadas en el rango muestral o en la desviación estándar muestral se usan para controlar la variabilidad del proceso. Muchas características de la calidad no se miden en una escala continua o siquiera cuantitativa. En estos casos, cada unidad de un producto puede juzgarse como conformable o no conformable con base en si posee o no ciertos atributos, o puede contarse el número de inconformidades (defectos) que aparecen en una unidad del producto. Las cartas de control para tales características de la calidad se llaman **cartas de control de atributos**.

Las cartas de control han tenido una larga historia de uso en la industria. Hay al menos cinco razones de su aceptación.

1. **Las cartas de control son una técnica probada para mejorar la productividad**. Un programa de cartas de control exitoso reducirá las piezas de desecho y las tareas de reprocesamiento, que son los principales depredadores de la productividad en *cualquier* operación. Cuando se reducen las piezas de desecho y las tareas de reprocesamiento, aumenta la productividad, se reducen los costos y se incrementa la capacidad de producción (medida en el número de piezas *aceptables* por hora).

2. **Las cartas de control son un recurso eficaz para prevenir defectos**. La carta de control ayuda a mantener el proceso bajo control, lo cual es consistente con la filosofía "hazlo bien a la primera". Nunca es más barato separar las unidades "buenas" de las "malas" en una fase posterior de lo que es hacerlas correctamente desde el principio. Si no se cuenta con un control del proceso eficaz, se estará pagando a alguien por hacer un producto no conformable.

3. **Las cartas de control previenen los ajustes innecesarios del proceso**. Una carta de control puede distinguir entre el ruido de fondo y una variación anormal; ningún otro recurso, incluyendo un operador humano, es tan eficaz para hacer esta distinción. Si los operadores del proceso ajustan el proceso con base en pruebas periódicas sin relación con un programa de carta de control, con frecuencia tendrán una reacción exagerada ante el ruido de fondo y realizarán ajustes innecesarios. Estos ajustes innecesarios por lo general redundan en un deterioro del desempeño del proceso. En otras palabras, la carta de control es consecuente con la filosofía "si no está descompuesto, no le metas mano".

4. **Las cartas de control proporcionan información de diagnóstico**. Frecuentemente, el patrón de los puntos de una carta de control incluirá información que tiene un valor de diagnóstico para un operador o ingeniero experimentado. Esta información permite al operador implementar un cambio en el proceso que mejorará su desempeño.

5. **Las cartas de control proporcionan información acerca de la capacidad del proceso.** La carta de control ofrece información acerca del valor de parámetros importantes del proceso y de su estabilidad con el tiempo. Esto permite hacer una estimación de la capacidad del proceso. Esta información es de gran utilidad para los diseñadores de productos y procesos.

Las cartas de control se encuentran entre las herramientas de control administrativo más efectivas, y son tan importantes como el control de costos y el control de materiales. La tecnología moderna de las computadoras ha facilitado la implementación de las cartas de control en cualquier tipo de proceso, pues la recolección y el análisis de datos puede llevarse a cabo en una microcomputadora o en la terminal de una red de área local en tiempo real, en línea dentro del centro de trabajo.

15-4.2 Diseño de una carta de control

A fin de ilustrar los conceptos anteriores, se presenta un ejemplo simplificado de una carta de control. En la fabricación de anillos para pistones de motores de automóviles, el diámetro interior de los anillos es una característica crítica de la calidad. La media del diámetro interior del proceso es 74 mm, y se sabe que la desviación estándar del diámetro de los anillos es 0.01 mm. En la figura 15-3 se muestra una carta de control del diámetro promedio de los anillos. Cada hora se toma una muestra aleatoria de cinco anillos, se calcula el diámetro promedio de los anillos de la muestra (por ejemplo, $\bar{x}$), y $\bar{x}$ se grafica en la carta. Debido a que esta carta de control utiliza la media muestral $\bar{X}$ para monitorear la media del proceso, suele llamársele carta de control $\bar{X}$. Obsérvese que todos los puntos se localizan dentro de los límites de control, por lo que la carta indica que el proceso está bajo control estadístico.

Considérese ahora cómo se determinaron los límites de control. El promedio del proceso es 74 mm, y la desviación estándar del proceso es $\sigma = 0.01$ mm. Entonces, si se toman muestras de tamaño $n = 5$, la desviación estándar del promedio muestral $\bar{X}$ es

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{5}} = 0.0045$$

Por lo tanto, si el procesto está bajo control con un diámetro promedio de 74 mm, al utilizar el teorema del límite central para suponer que $\bar{X}$ tiene una distribución aproximadamente normal,

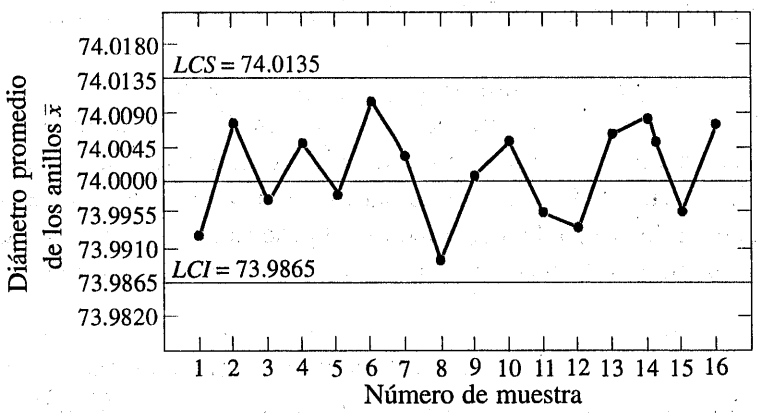


Figura 15-3 Carta de control $\bar{X}$ del diámetro de los anillos para pistones.

se esperaría que aproximadamente $100(1 - \alpha)$ por ciento de los diámetros muestrales promedio X se localizarán entre $74 + z_{\alpha/2}(0.0045)$ y $74 - z_{\alpha/2}(0.0045)$. Como se señaló arriba, se acostumbra elegir que la constante $z_{\alpha/2}$ sea 3, de donde los límites de control superior e inferior son

$$LCS = 74 + 3(0.0045) = 74.0135$$

y

$$LCI = 74 - 3(0.0045) = 73.9865$$

como se muestra en la carta de control. Se trata de los límites de control 3-sigma a los que se hizo referencia arriba. Obsérvese que el uso de límites 3-sigma implica que $\alpha = 0.0027$; es decir, la probabilidad de que los puntos graficados se localicen fuera de los límites de control cuando el proceso está bajo control es 0.0027. La anchura de los límites de control tiene una relación inversa con el tamaño de la muestra n para un múltiplo dado de sigma. Elija los límites de control es equivalente a establecer la región crítica para probar la hipótesis

$$H_0 : \mu = 74$$
$$H_1 : \mu \neq 74$$

donde $\sigma = 0.01$ es conocido. En esencia, la carta de control prueba esta hipótesis varias veces en diferentes puntos del tiempo.

Al diseñar una carta de control, debe especificarse tanto el tamaño de la muestra que va a usarse como la frecuencia del muestreo. En general, las muestras más grandes facilitarán la detección de los corrimientos pequeños en el proceso. Cuando se elige el tamaño de la muestra, deberá tenerse presente el tamaño del corrimientos que está tratando de detectarse. Si quiere detectarse un corrimiento relativamente grande del proceso, entonces se usan tamaños de las muestras más pequeños que los que se emplearían si el corrimiento de interés fuera relativamente pequeño.

También es necesario determinar la frecuencia del muestreo. La situación más deseable desde el punto de vista de la detección de corrimientos sería tomar muestras grandes con mucha frecuencia; sin embargo, esto por lo general no resulta económicamente factible. Se trata del problema general de *distribución del esfuerzo de muestreo.* Es decir, si se tomarán muestras pequeñas en intervalos cortos o muestras más grandes en intervalos más largos. La práctica actual en la industria tiende a favorecer las muestras más pequeñas y más frecuentes, en particular en los procesos de manufactura con un alto volumen de producción, o donde pueden ocurrir diversas clases de causas asignables. Por otra parte, a medida que se desarrollan las tecnologías de sensores y mediciones automatizadas, se está haciendo posible incrementar considerablemente la frecuencia del muestreo. En última instancia, cada unidad puede probarse cuando se está fabricando. Esta capacidad no eliminará la necesidad de las cartas de control, ya que el sistema de prueba no corregirá los defectos. El número mayor de datos aumentará la eficiencia del control del proceso y mejorará la calidad.

15-4.3 Subgrupos racionales

Una idea fundamental en el uso de las cartas de control es recabar los datos muestrales de acuerdo con lo que Shewhart denominaba el concepto del **subgrupo racional**. En general, esto

significa que los subgrupos o muestras deberán seleccionarse de tal modo que, en la medida de lo posible, la variabilidad de las observaciones dentro de un subgrupo deberá incluir toda la variabilidad fortuita o natural y excluir la variabilidad asignable. Entonces, los límites de control representarán cotas para toda la variabilidad fortuita y no para la variabilidad asignable. Por consiguiente, las causas asignables tenderán a generar puntos que se localizan fuera de los límites de control, en tanto que la variabilidad fortuita tenderá a generar puntos que se localizan dentro de los límites de control.

Cuando las cartas de control se aplican a procesos de producción, el orden en el tiempo de la producción es una base lógica para el establecimiento de subgrupos racionales. Sin embargo, incluso cuando se preserva el orden en el tiempo, sigue existiendo la posibilidad de formar subgrupos erróneamente. Si algunas de las observaciones del subgrupo se toman al final de un turno de 8 horas y las observaciones restantes se toman al principio del siguiente turno de 8 horas, entonces quizá no se detecte alguna diferencia entre los turnos. El orden en el tiempo frecuentemente es una base apropiada para formar subgrupos, debido a que permite detectar causas asignables que ocurren en el tiempo.

Se emplean dos enfoques generales para construir subgrupos racionales. En el primer enfoque, cada subgrupo se compone de unidades que se fabricaron al mismo tiempo (o tan cercanamente como sea posible). Este enfoque se utiliza cuando la finalidad principal de la carta de control es detectar corrimientos en el proceso. Minimiza la variabilidad debida a causas asignables *dentro* de una muestra, y maximiza la variabilidad *entre* las muestras en caso de que haya causas asignables presentes. Proporciona asimismo mejores estimaciones de la desviación estándar del proceso cuando se emplean cartas de control de variables. Este enfoque para la formación de subgrupos racionales proporciona en esencia una "instantánea" del proceso en cada punto del tiempo en que se recolecta una muestra.

En el segundo enfoque, cada muestra se compone de unidades del producto que son representativas de *todas* las unidades que se han producido desde que se tomó la última muestra. En esencia, cada subgrupo es una *muestra aleatoria* de la producción *total* del proceso en el intervalo del muestreo. Este método para formar los subgrupos racionales suele usarse cuando la carta de control se emplea para tomar decisiones relativas a la aceptación de todas las unidades del producto que se han producido desde la última muestra. De hecho, si el proceso se corre a un estado fuera de control y después vuelve a estar bajo control *entre* las muestras, en ocasiones se argumenta que el primer método para formar los subgrupos racionales definido arriba resultará ineficaz contra este tipo de corrimientos y que, por tanto, debe usarse el segundo método.

Cuando el subgrupo racional es una muestra aleatoria de todas las unidades producidas en el intervalo del muestreo, deberá tenerse especial cuidado al interpretar las cartas de control. Si la media del proceso oscila entre varios niveles durante el intervalo entre las muestras, entonces el rango de las observaciones dentro de la muestra puede ser en consecuencia relativamente grande. Es la variabilidad dentro de la muestra la que determina la anchura de los límites de control en una carta $\overline{X}$, por lo que esta práctica producirá límites con una anchura mayor en la carta $\overline{X}$. Esto hace más difícil detectar los corrimientos en la media. De hecho, con frecuencia es posible hacer que *cualquier* proceso parezca estar bajo control estadístico simplemente ampliando el intervalo entre las observaciones de la muestra. Asimismo, en el caso de las cartas de control basadas en el rango o la desviación estándar es posible que corrimientos en el promedio del proceso produzcan puntos que se localizan fuera de control, aun cuando no haya tenido lugar ningún corrimiento en la variabilidad del proceso.

Hay otras bases para formar los subgrupos racionales. Por ejemplo, suponga que un proceso consta de varias máquinas cuya producción se canaliza en un flujo común. Si el muestreo se hace en este flujo común de la producción, será muy difícil detectar si alguna de las máquinas está fuera de control o no. Un enfoque lógico para la formación de los subgrupos racionales en este caso es aplicar las técnicas de las cartas de control a la producción de cada máquina individual. En ocasiones será necesario aplicar este concepto a diferentes secciones de la misma máquina, a diferentes estaciones de trabajo, a diferentes operadores, etcétera.

El concepto de subgrupo racional es de suma importancia. La selección apropiada de las muestras requiere la consideración atenta del proceso, a fin de conseguir la mayor cantidad de información útil posible del análisis de la carta de control.

15-4.4 Análisis de patrones en cartas de control

Una carta de control puede indicar una condición fuera de control, ya sea cuando uno o más puntos se localizan fuera de los límites de control o bien cuando los puntos graficados presentan algún patrón de comportamiento no aleatorio. Por ejemplo, considérese la carta $\bar{X}$ que se muestra en la figura 15-4. Aun cuando los 25 puntos se localizan dentro de los límites de control, los puntos no indican un estado de control estadístico, debido a que la apariencia del patrón que muestran es en apariencia no aleatoria. Específicamente, se observa que 19 de los 25 puntos se localizan abajo de la línea central, mientras que sólo 6 de ellos se localizan arriba de la línea central. Si los puntos fueran realmente aleatorios, se esperaría una distribución más pareja de los mismos arriba y abajo de la línea central. Se observa asimismo que después del cuarto punto, cinco puntos consecutivos se incrementan en magnitud. A este arreglo de puntos se le denomina una **corrida**. Puesto que las observaciones son crecientes, podría llamársele una corrida ascendente; de manera similar, a una secuencia de puntos decrecientes se le llama corrida descendente. Esta carta de control tiene una serie ascendente inusualmente larga (empezando en el cuarto punto) y una serie descendente inusualmente larga (a partir del punto dieciocho).

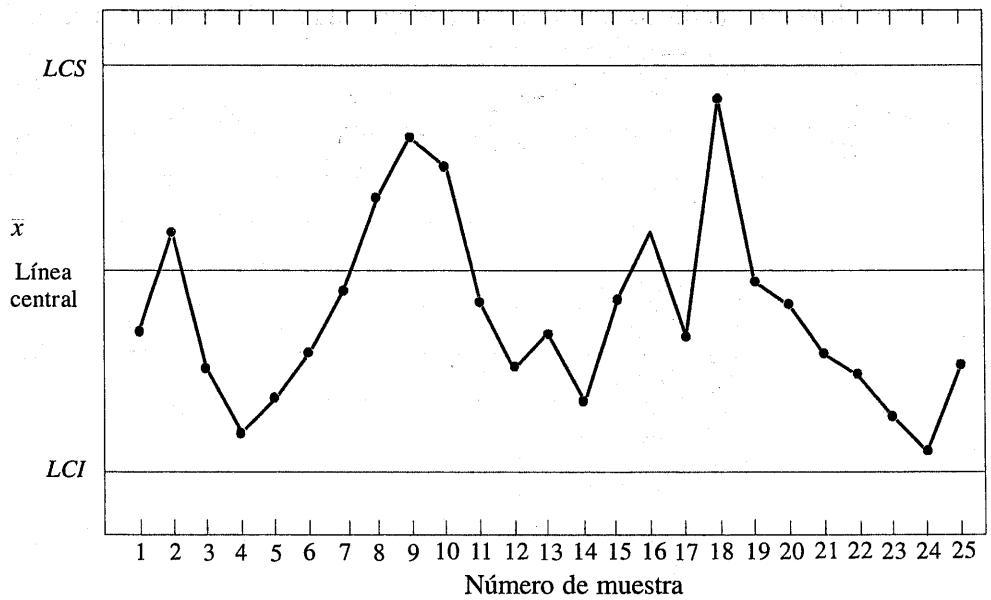


Figura 15-4 Una carta de control $\bar{X}$.

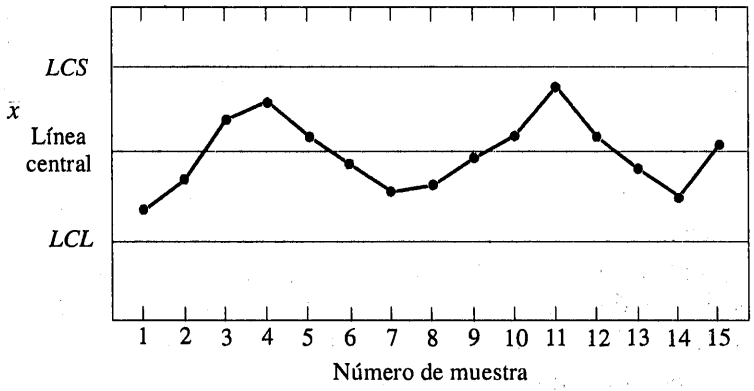


Figura 15-5 Una carta $\bar{X}$ con un patrón cíclico.

En general, una corrida se define como una secuencia de observaciones del mismo tipo. Además de las corridas ascendentes y descendentes, las observaciones podrían tipificarse como las que se localizan arriba y abajo de la línea central, respectivamente, de tal suerte que dos puntos consecutivos arriba de la línea central serían una corrida de longitud 2.

La probabilidad de que en una muestra de puntos aleatorios ocurra una corrida con una longitud de 8 o más puntos es muy reducida. Por consiguiente, cualquier tipo de corrida de longitud ocho o más con frecuencia se toma como una señal de una condición fuera de control. Por ejemplo, 8 puntos consecutivos a uno de los lados de la línea central indicará que el proceso está fuera de control.

Aun cuando las corridas son una medida importante del comportamiento no aleatorio en una carta de control, otros tipos de patrones también pueden indicar una condición fuera de control. Por ejemplo, considérese la carta $\bar{X}$ de la figura 15-5. Obsérvese que los promedios muestrales graficados muestran un comportamiento cíclico, no obstante que todos están dentro de los límites de control. Este patrón puede indicar un problema con el proceso, tal como fatiga del operador, reaprovisionamiento de materia prima, o aumento de la temperatura o del esfuerzo. El rendimiento puede mejorarse eliminando o reduciendo las fuentes de variabilidad que

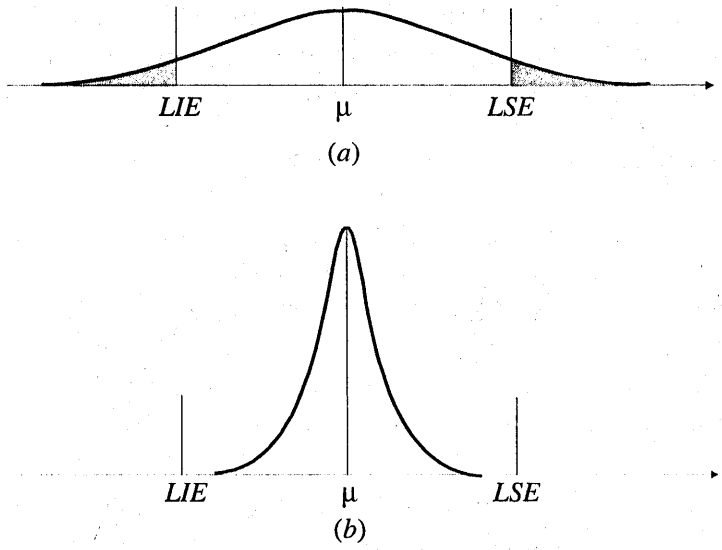


Figura 15-6 ***a***) Variabilidad con el patrón cíclico. ***b***) Variabilidad con el patrón cíclico eliminado.

ocasionan este comportamiento cíclico (véase la figura 15-6). En la figura 15-6, *LIE* y *LSE* denotan los límites inferior y superior de las especificaciones del proceso. Estos límites representan cotas dentro de las cuales deben localizarse los productos aceptables y con frecuencia se basan en requerimientos del cliente.

Se trata de un problema de **identificación de patrones**, es decir, de identificar los patrones sistemáticos o no aleatorios en la carta de control y de determinar la razón de este comportamiento. La habilidad para interpretar un patrón particular en términos de causas asignables requiere experiencia y conocimiento del proceso. Es decir, no basta con conocer los principios estadísticos de las cartas de control, sino que también es necesario contar con una comprensión detallada del proceso.

El *Western Electric Handbook* sugiere un conjunto de reglas de decisión para detectar patrones no aleatorios en cartas de control. Específicamente, las **reglas Western Electric** concluirían que el proceso está fuera de control si se cumple cualquiera de las condiciones siguientes

1. Uno de los puntos se localiza fuera de los límites de control 3-sigma.
2. Dos de tres puntos consecutivos se localizan después de un límite 2-sigma.
3. Cuatro de cinco puntos consecutivos se localizan a una distancia de 1-sigma o más de la línea central.
4. Ocho puntos consecutivos se localizan a uno de los lados de la línea central.

Los autores han encontrado que en la práctica estas reglas son muy efectivas para aumentar la sensibilidad de las cartas de control. Las reglas 2 y 3 se aplican a uno de los lados de la línea central a la vez. Es decir, un punto arriba del límite 2-sigma *superior* seguido inmediatamente por un punto abajo del límite 2-sigma *inferior* no indicaría una alarma de fuera de control.

En la figura 15-7 se muestra una carta de control $\overline{X}$ para el proceso de los anillos para pistones con los límites 1-sigma, 2-sigma y 3-sigma usados en el procedimiento Western Electric. Obsérvese que estos límites interiores (en ocasiones llamados **límites de advertencia**) dividen la carta de control en tres zonas A, B y C a cada lado de la línea central. Por consiguiente, las

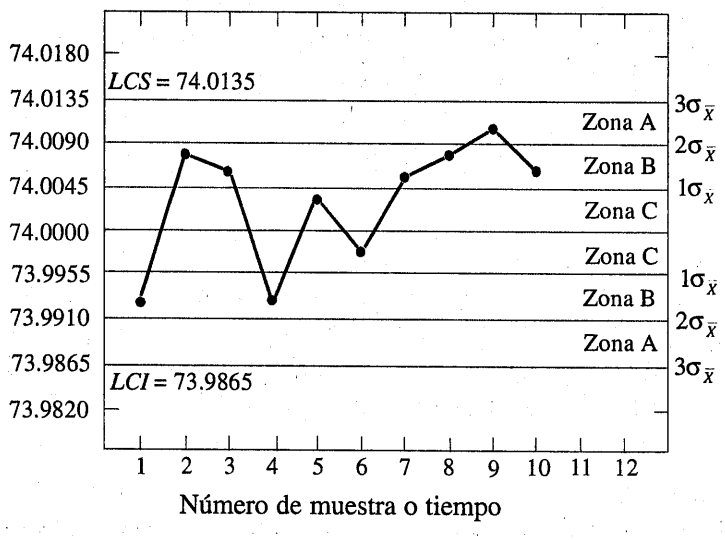


Figura 15-7 Las reglas Western Electric de zonas.

reglas Western Electric se llaman en ocasiones las **reglas de zona** para cartas de control. Obsérvese que los cuatro últimos puntos se localizan en la zona B o después de ella. Por tanto, puesto que cuatro de cinco puntos consecutivos exceden el límite 1-sigma, con el procedimiento Western Electric se concluirá que el patrón es no aleatorio y que el proceso está fuera de control.

15-5 CARTAS DE CONTROL $\bar{X}$ Y R

Cuando se trabaja con una característica de la calidad que puede expresarse como una medición, se acostumbra monitorear tanto el valor medio de la característica de la calidad como su variabilidad. El control sobre la calidad promedio se ejerce por la carta de control de los promedios, conocida generalmente como carta $\bar{X}$. La variabilidad del proceso puede controlarse con una carta del rango (carta R) o bien con una carta de la desviación estándar (carta S), dependiendo de la manera en que se estime la desviación estándar de la población. Se revisará únicamente la carta R.

Suponga que se conocen la media μ y la desviación estándar σ del proceso y que la característica de la calidad tiene una distribución normal. Considérese la carta $\bar{X}$. Como se mencionó anteriormente, puede usarse μ como la línea central de la carta de control, y los límites 3-sigma superior e inferior pueden colocarse en $LCS = \mu + 3\sigma/\sqrt{n}$ y $LCI = \mu - 3\sigma/\sqrt{n}$, respectivamente.

Cuando no se conocen los parámetros μ y σ, por lo general se estiman con base en muestras preliminares, las cuales se toman cuando se considera que el proceso está bajo control. Se recomienda el uso de por lo menos 20 a 25 muestras preliminares. Suponga que se cuenta con m muestras preliminares, cada una de tamaño n. De manera típica, n será 4, 5 o 6; estos tamaños de las muestras relativamente pequeños se usan ampliamente y con frecuencia surgen de la construcción de subgrupos racionales. Sea $\bar{X}_i$ la media muestral de la i-ésima muestra. Entonces la media de la población, μ, se estima por medio de la *gran media*

$$\bar{\bar{X}} = \frac{1}{m}\sum_{i=1}^{m}\bar{X}_i \qquad (15\text{-}2)$$

Por tanto, $\bar{\bar{X}}$ puede tomarse como la línea central en la carta de control $\bar{X}$.

Puede estimarse σ ya sea a partir de la desviación estándar o bien del rango de las observaciones dentro de cada muestra. Dado que se utiliza con mayor frecuencia en la práctica, el análisis presente se restringe al método del rango. El tamaño de la muestra es relativamente pequeño, por lo que la pérdida de eficiencia al estimar σ a partir de los rangos muestrales es reducida.

Se necesita la relación entre el rango R de una muestra de una población normal con parámetros conocidos y la desviación estándar de esa población. Puesto que R es una variable aleatoria, la cantidad $W = R/\sigma$, llamada el rango relativo, también es una variable aleatoria. Los parámetros de la distribución de W se han determinado para cualquier tamaño de la muestra n.

A la media de la distribución de W se le llama d_2, y en la tabla VI del apéndice se presenta una tabla de d_2 para varios valores de n. A la desviación estándar de W se le llama d_3. Debido a que $R = \sigma W$

$$\mu_R = d_2\sigma$$
$$\sigma_R = d_3\sigma$$

Sea R_i el rango de la i-ésima muestra, y sea

$$\bar{R} = \frac{1}{m}\sum_{i=1}^{m} R_i \tag{15-3}$$

el rango promedio. Entonces $\bar{R}$ es un estimador de μ_R, y un estimador de σ es

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \tag{15-4}$$

Por lo tanto, como límites de control superior e inferior de la carta $\bar{X}$ pueden usarse

$$LCS = \bar{X} + \frac{3}{d_2\sqrt{n}}\bar{R}$$
$$LCI = \bar{\bar{X}} - \frac{3}{d_2\sqrt{n}}\bar{R} \tag{15-5}$$

Se define la constante

$$A_2 = \frac{3}{d_2\sqrt{n}} \tag{15-6}$$

Entonces, una vez que se han calculado los valores muestrales $\bar{\bar{x}}$ y $\bar{r}$, los parámetros de la carta de control $\bar{X}$ pueden definirse de la siguiente manera:

Definición

La línea central y los límites de control superior e inferior de una carta de control $\bar{X}$ son

$$LCS = \bar{\bar{x}} + A_2\bar{r}$$
$$LC = \bar{\bar{x}} \tag{15-7}$$
$$LCI = \bar{\bar{x}} - A_2\bar{r}$$

donde la constante A_2 se tabula para varios tamaños de las muestras en la tabla VI del apéndice.

También es sencillo determinar los parámetros de la carta R. Obviamente, la línea central será $\bar{R}$. Para determinar los límites de control, se necesita una estimación de σ_R, la desviación estándar de R. De nueva cuenta, suponiendo que el proceso está bajo control, la distribución del rango relativo, W, será de utilidad. Puesto que no se conoce σ, σ_R puede estimarse como

$$\hat{\sigma}_R = d_3\hat{\sigma} = d_3 \frac{R}{d_2}$$

y como límites de control superior e inferior en la carta R se usarían

$$LCS = \bar{R} + \frac{3d_3}{d_2}\bar{R} = \left(1 + \frac{3d_3}{d_2}\right)\bar{R}$$

$$LCI = \bar{R} - \frac{3d_3}{d_2}\bar{R} = \left(1 - \frac{3d_3}{d_2}\right)\bar{R}$$

Al hacer $D_3 = 1 - 3d_3/d_2$ y $D_4 = 1 + 3d_3/d_2$ se llega a la siguiente definición.

Definición

La línea central y los límites de control superior e inferior de una carta R son

$$\begin{aligned} LCS &= D_4\bar{r} \\ LC &= \bar{r} \\ LCI &= D_3\bar{r} \end{aligned} \tag{15-8}$$

donde $\bar{r}$ es el rango muestral promedio, y las constantes D_3 y D_4 se tabulan para varios tamaños de las muestras en la tabla XII del apéndice.

Cuando se usan muestras preliminares para construir los límites de una carta de control, se acostumbra tratar estos límites como valores de prueba. Por lo tanto, las m medias y rangos muestrales deberán graficarse en las cartas apropiadas, y habrán que investigarse cualesquiera puntos que excedan los límites de control. Si se descubren las causas asignables de estos puntos, deberán eliminarse y determinarse los nuevos límites de las cartas de control. De este modo, el proceso eventualmente se llevará a un estado de control estadístico y podrán determinarse sus capacidades inherentes. Entonces pueden contemplarse otros cambios en el centrado y la dispersión del proceso. Asimismo, con frecuencia se estudia primero la carta R porque si la variabilidad del proceso no es constante con el tiempo, los límites de control calculados para la carta $\bar{X}$ pueden llevar a conclusiones falsas.

EJEMPLO 15-1

La pieza de una turbina de avión se fabrica mediante un proceso de colada de precisión. La abertura del aspa en este vaciado es un parámetro funcional importante de la pieza. Se ilustrará el uso de las cartas de control $\bar{X}$ y R para determinar la estabilidad estadística de este proceso. En la tabla 15-1 se presentan 20 muestras de cinco piezas cada una. Los valores dados en la tabla se han codificado usando los tres últimos dígitos de la dimensión; es decir, 31.6 debería ser 0.50316 pulgadas.

Las cantidades $\bar{\bar{x}} = 33.3$ y $\bar{r} = 5.8$ se muestran al final de la tabla 15-1. El valor de A_2 para muestras de tamaño 5 es $A_2 = 0.577$. Entonces los límites de control de prueba de la carta $\bar{X}$ son

$$\bar{\bar{x}} = \pm A_2\bar{r} = 3.32 \pm (0.577)(5.8) = 33.32 \pm 3.35$$

o bien

$$LCS = 36.67$$

$$LCI = 29.97$$

Tabla 15-1 Mediciones de la abertura de las aspas

Número de muestra	x_1	x_2	x_3	x_4	x_5	$\bar{x}$	r
1	33	29	31	32	33	31.6	4
2	33	31	35	37	31	33.4	6
3	35	37	33	34	36	35.0	4
4	30	31	33	34	33	32.2	4
5	33	34	35	33	34	33.8	2
6	38	37	39	40	38	38.4	3
7	30	31	32	34	31	31.6	4
8	29	39	38	39	39	36.8	10
9	28	33	35	36	43	35.0	15
10	38	33	32	35	32	34.0	6
11	28	30	28	32	31	29.8	4
12	31	35	35	35	34	34.0	4
13	27	32	34	35	37	33.0	10
14	33	33	35	37	36	34.8	4
15	35	37	32	35	39	35.6	7
16	33	33	27	31	30	30.8	6
17	35	34	34	30	32	33.0	5
18	32	33	30	30	33	31.6	3
19	25	27	34	27	28	28.2	9
20	35	35	36	33	30	33.8	6
						$\bar{\bar{x}} = 33.32$	$\bar{r} = 5.8$

Para la carta R, los límites de control de prueba son

$$LCS = D_4\bar{r} = (2.115)(5.8) = 12.27$$
$$LCI = D_3\bar{r} = (0)(5.8) = 0$$

En la figura 15-8 se muestran las cartas de control $\bar{X}$ y R con estos límites de control de prueba. Obsérvese que las muestras 6, 8, 11 y 19 están fuera de control en la carta $\bar{X}$ y que la muestra 9 está fuera de control en la carta R. (Estos puntos se marcan con un "1" porque infringen la primera regla Western Electric.) Suponga que todas estas causas asignables pueden rastrearse hasta una herramienta defectuosa en el área del moldeo de la cera. Deberán descartarse estas cinco muestras y calcular de nuevo los límites de las cartas $\bar{X}$ y R. Los nuevos límites revisados son, para la carta $\bar{X}$,

$$LCS = \bar{\bar{x}} + A_2\bar{r} = 33.21 + (0.577)(5.0) = 36.10$$
$$LCI = \bar{\bar{x}} - A_2\bar{r} = 33.21 - (0.577)(5.0) = 30.33$$

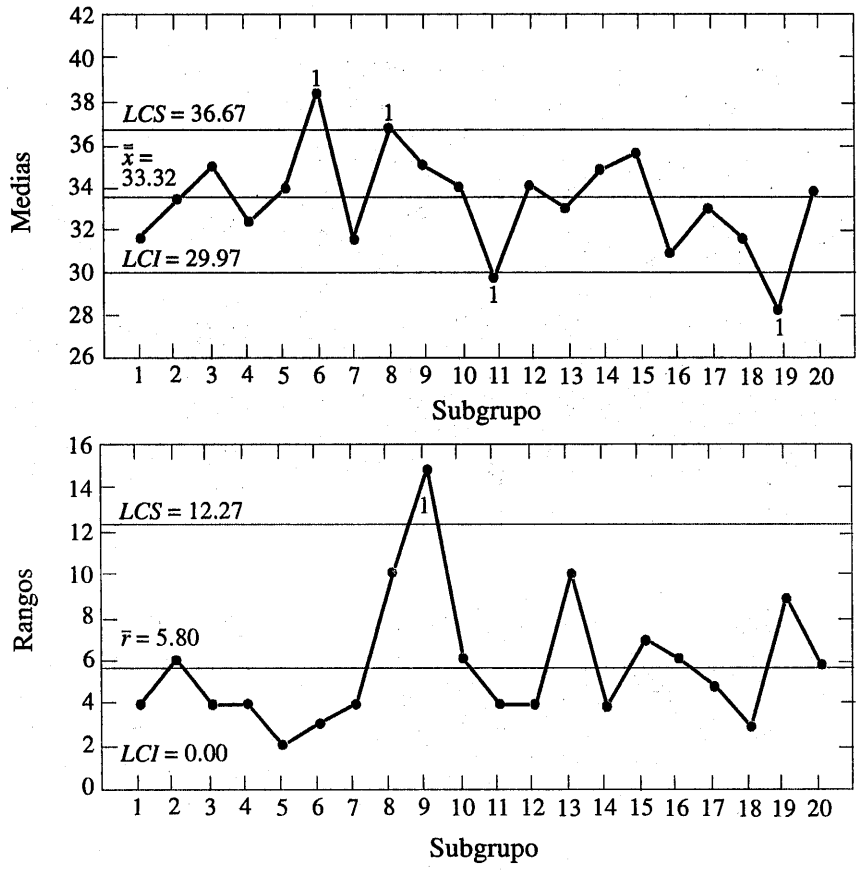


Figura 15-8 Las carta de control $\bar{X}$ y R para la abertura de las aspas.

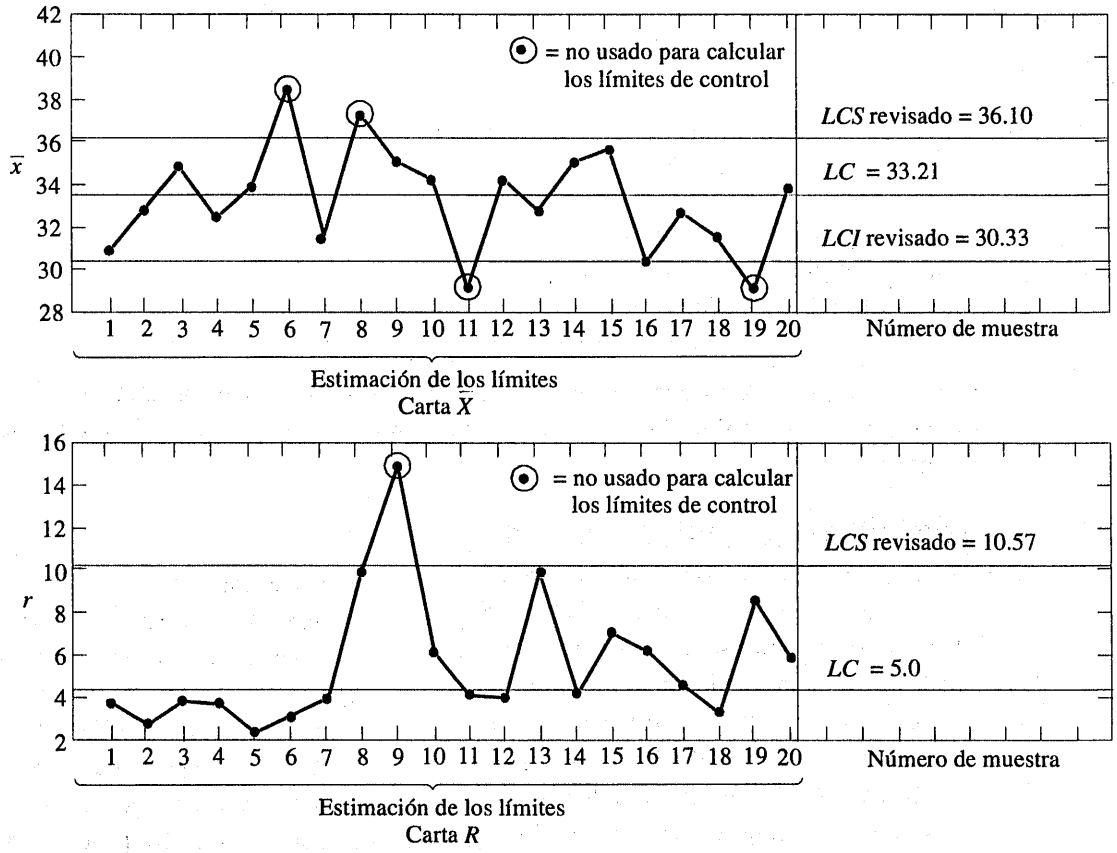


Figura 15-9 Las carta de control $\bar{X}$ y R para la abertura de las aspas, límites revisados.

y para la carta R,

$$LCS = D_4\bar{r} = (2.115)(5.0) = 10.57$$
$$LCI = D_3\bar{r} = (0)(5.0) = 0$$

En la figura 15-9 se muestran las cartas de control revisadas. Obsérvese que las 20 primeras muestras preliminares se han tratado como **datos de estimación** con los cuales se establecen los límites de control. Estos límites pueden usarse ahora para juzgar el control estadístico de la producción futura. Conforme se vaya teniendo cada nueva muestra, los valores de $\bar{x}$ y r deberán calcularse y graficarse en las cartas de control. Puede ser conveniente revisar los límites periódicamente, incluso si el proceso se mantiene estable. Los límites deberán revisarse siempre que se hagan mejoras en el proceso.

Construcción de las cartas de control $\bar{X}$ y R con computadora

Muchos programas de computadora crean cartas de control $\bar{X}$ y R. En las figuras 15-8 y 15-9 se muestran cartas similares a las que produce el programa Minitab para los datos de las aberturas de las aspas del ejemplo 15-1. Este programa permitirá al usuario seleccionar cualquier múltiplo de sigma como la anchura de los límites de control así como utilizar las reglas Western

Tabla 15-2 Reporte resumido del programa Minitab para los datos de la abertura de las aspas del ejemplo 15-1

Resultado de la prueba para la carta X barra
PRUEBA 1. Un punto a más de 3.00 sigmas de la línea central
La prueba falló en los puntos: 6 8 11 19

Resultado de la prueba para la carta R
PRUEBA 1. Un punto a más de 3.00 sigmas de la línea central
La prueba falló en los puntos: 9

Electric para detectar los puntos que están fuera de control. El programa también prepara un reporte resumido como el de la tabla 15-2, excluyendo los subgrupos del cálculo de los límites de control.

EJERCICIOS DE LA SECCIÓN 15-5

15-1. Se usa una matriz de extrusión para producir varillas de aluminio. El diámetro de las varillas es una característica crítica de la calidad. En la tabla siguiente se muestran los valores $\bar{x}$ y r para 20 muestras de cinco varillas cada una. Las especificaciones de las varillas son 0.5035 ± 0.0010 pulgadas. Los valores que se presentan son los tres últimos dígitos de la medición; es decir, 34.2 se lee como 0.50342.

Muestra	$\bar{x}$	r
1	34.2	3
2	31.6	4
3	31.8	4
4	33.4	5
5	35.0	4
6	32.1	2
7	32.6	7
8	33.8	9
9	34.8	10
10	38.6	4
11	35.4	8
12	34.0	6
13	36.0	4
14	37.2	7
15	35.2	3
16	33.4	10
17	35.0	4
18	34.4	7
19	33.9	8
20	34.0	4

a) Utilizando todos los datos, encuentre los límites de control de prueba de las cartas $\bar{X}$ y R, construya las cartas y grafique los datos.
b) Use los límites de control de prueba del inciso a) para identificar los puntos fuera de control. De ser necesario, revise los límites de control, suponiendo que cualquier muestra que se localice fuera de los límites de control puede eliminarse.

15-2. Se sacan 25 muestras de tamaño 5 de un proceso en intervalos de una hora, y se obtienen los siguientes datos:

$$\sum_{i=1}^{25} \bar{x}_i = 362.75 \qquad \sum_{i=1}^{25} r_i = 8.60$$

Encuentre los límites de control de prueba de las cartas $\bar{X}$ y R.

15-3. Se monitorea la longitud total del cuerpo de un encendedor de cigarrillos usado en una aplicación automotriz utilizando carta $\bar{X}$ y R. En la tabla siguiente se muestra la longitud de 20 muestras de tamaño 4. (Las mediciones están codificadas a partir de 5.00 mm; es decir, 15 es 5.15 mm.)

a) Utilizando todos los datos, encuentre los límites de control de prueba de las cartas $\bar{X}$ y R, construya las cartas y grafique los datos.
b) Use los límites de control de prueba del inciso a) para identificar los puntos fuera de control. De ser necesario, revise los lí-

Muestra	Observación 1	2	3	4
1	15	10	8	9
2	14	14	10	6
3	9	10	9	11
4	8	6	9	13
5	14	8	9	12
6	9	10	7	13
7	15	10	12	12
8	14	16	11	10
9	11	7	16	10
10	11	14	11	12
11	13	8	9	5
12	10	15	8	10
13	8	12	14	9
14	15	12	14	6
15	13	16	9	5
16	14	8	8	12
17	8	10	16	9
18	8	14	10	9
19	13	15	10	8
20	9	7	15	8

mites de control, suponiendo que cualquier muestra que se localice fuera de los límites de control puede eliminarse.

15-4. Cada hora se recolectan muestras de tamaño $n = 6$ de un proceso. Después de que se han recolectado 20 muestras, se calcula $\bar{\bar{x}} = 20.0$ y $\bar{r}/d_2 = 1.4$. Encuentre los límites de control de prueba de las cartas $\bar{X}$ y R.

15-5. Deben construirse las cartas de control $\bar{X}$ y R para una importante característica de la calidad. El tamaño de la muestra es $n = 5$, y $\bar{x}$ y r se calculan para cada una de 35 muestras preliminares. Los datos resumidos son

$$\sum_{i=1}^{35} \bar{x}_i = 7\,805 \qquad \sum_{i=1}^{35} r_i = 1\,200$$

a) Encuentre los límites de control de prueba de las cartas $\bar{X}$ y R.

b) Suponiendo que el proceso está bajo control, estime la media y la desviación estándar del proceso.

15-6. El espesor de una pieza metálica es un parámetro importante de la calidad. En la tabla siguiente se presentan los datos del espesor (en pulgadas) de 25 muestras de cinco piezas cada una.

Número de muestra	x_1	x_2	x_3	x_4	x_5
1	0.0629	0.0636	0.0640	0.0635	0.0640
2	0.0630	0.0631	0.0622	0.0625	0.0627
3	0.0628	0.0631	0.0633	0.0633	0.0630
4	0.0634	0.0630	0.0631	0.0632	0.0633
5	0.0619	0.0628	0.0630	0.0619	0.0625
6	0.0613	0.0629	0.0634	0.0625	0.0628
7	0.0630	0.0639	0.0625	0.0629	0.0627
8	0.0628	0.0627	0.0622	0.0625	0.0627
9	0.0623	0.0626	0.0633	0.0630	0.0624
10	0.0631	0.0631	0.0633	0.0631	0.0630
11	0.0635	0.0630	0.0638	0.0635	0.0633
12	0.0623	0.0630	0.0630	0.0627	0.0629
13	0.0635	0.0631	0.0630	0.0630	0.0630
14	0.0645	0.0640	0.0631	0.0640	0.0642
15	0.0619	0.0644	0.0632	0.0622	0.0635
16	0.0631	0.0627	0.0630	0.0628	0.0629
17	0.0616	0.0623	0.0631	0.0620	0.0625
18	0.0630	0.0630	0.0626	0.0629	0.0628
19	0.0636	0.0631	0.0629	0.0635	0.0634
20	0.0640	0.0635	0.0629	0.0635	0.0634
21	0.0628	0.0625	0.0616	0.0620	0.0623
22	0.0615	0.0625	0.0619	0.0619	0.0622
23	0.0630	0.0632	0.0630	0.0631	0.0630
24	0.0635	0.0629	0.0635	0.0631	0.0633
25	0.0623	0.0629	0.0630	0.0626	0.0628

a) Utilizando todos los datos, encuentre los límites de control de prueba de las cartas $\bar{X}$ y R, construya las cartas y grafique los datos. ¿El proceso está bajo control estadístico?

b) Use los límites de control de prueba del inciso a) para identificar los puntos fuera de control. Enliste los números de las muestras de los puntos fuera de control.

15-7. El contenido de cobre de un baño de plateado se mide tres veces por día, y los resultados se reportan en ppm. En la tabla siguiente se muestran los valores de $\bar{x}$ y r para 25 días.

a) Utilizando todos los datos, encuentre los límites de control de prueba de las cartas $\bar{X}$ y R, construya las cartas y grafique los datos. ¿El proceso está bajo control estadístico?

b) De ser necesario, revise los límites de control calculados en el inciso a), suponiendo que cualquier muestra que se localice fuera de los límites de control puede eliminarse.

Día	$\bar{x}$	r	Día	$\bar{x}$	r
1	5.45	1.21	14	7.01	1.45
2	5.39	0.95	15	5.83	1.37
3	6.85	1.43	16	6.35	1.04
4	6.74	1.29	17	6.05	0.83
5	5.83	1.35	18	7.11	1.35
6	7.22	0.88	19	7.32	1.09
7	6.39	0.92	20	5.90	1.22
8	6.50	1.13	21	5.50	0.98
9	7.15	1.25	22	6.32	1.21
10	5.92	1.05	23	6.55	0.76
11	6.45	0.98	24	5.90	1.20
12	5.38	1.36	25	5.95	1.19
13	6.03	0.83			

15-6 CARTAS DE CONTROL PARA MEDICIONES INDIVIDUALES

En muchas situaciones, el tamaño de la muestra usado para el control del proceso es $n = 1$; es decir, la muestra consta de una unidad individual. Algunos ejemplos de estas situaciones son los siguientes.

1. Se usa la tecnología de inspección y medición automatizada, y se analiza cada unidad fabricada.
2. La rapidez con que se saca la producción es muy lenta, y no es conveniente permitir que se acumulen muestras de tamaño $n > 1$ antes de analizarlas.
3. Las mediciones repetidas del proceso tan sólo difieren por el error de laboratorio o análisis, como en muchos procesos químicos.
4. En las plantas de procesamiento, como en la fabricación de papel, las mediciones de algunos parámetros tales como el espesor del recubrimiento *a lo ancho* del rollo diferirá muy poco y producirá una desviación estándar que es por mucho demasiado pequeña si el objetivo es controlar el espesor del recubrimiento *a lo largo* del rollo.

En tales situaciones resulta conveniente la **carta de control para mediciones individuales**. En la carta de control para mediciones individuales se emplea el **rango móvil** de dos observaciones sucesivas para estimar la variabilidad del proceso. El rango móvil se define como $MR_i = |X_i - X_{i-1}|$.

Una estimación de σ es

$$\hat{\sigma} = \frac{\overline{MR}}{d_2} = \frac{\overline{MR}}{1.128}$$

dado que $d_2 = 1.128$ cuando se usan dos observaciones consecutivas para calcular un rango móvil. También es posible establecer una carta de control para el rango móvil usando D_3 y D_4 para $n = 2$. Los parámetros de estas cartas se definen de la siguiente manera.

Definición

La línea central y los límites de control superior e inferior de una carta de control para mediciones individuales son

$$LCS = \bar{x} + 3\frac{\overline{mr}}{d_2} = \bar{x} + 3\frac{\overline{mr}}{1.128}$$
$$LC = \bar{x} \qquad (15\text{-}9)$$
$$LCI = \bar{x} - 3\frac{\overline{mr}}{d_2} = \bar{x} - 3\frac{\overline{mr}}{1.128}$$

y para una carta de control para rangos móviles son

$$LCS = D_4\overline{mr} = 3.267\overline{mr}$$
$$LC = \overline{mr}$$
$$LCI = D_3\overline{mr} = 0$$

El procedimiento se ilustra en el ejemplo siguiente.

EJEMPLO 15-2

En la tabla 15-3 se muestran 20 observaciones de la concentración de la producción de un proceso químico. Las observaciones se hacen en intervalos de una hora. Si se hacen varias observaciones al mismo tiempo, la concentración observada diferirá tan sólo por el error de medición. Puesto que el error de medición es pequeño, sólo se hace una observación cada hora.

Para construir la carta de control para mediciones individuales, obsérvese que el promedio muestral de las 20 lecturas de la concentración es $\bar{x} = 99.1$ y que el promedio de los rangos móviles de dos observaciones mostrado en la última columna de la tabla 15-3 es $\overline{mr} = 2.59$. Para construir la carta para rangos móviles, se observa que $D_3 = 0$ y $D_4 = 3.267$ para $n = 2$. Por lo tanto, la carta para rangos móviles tiene la línea central $\overline{mr} = 2.59$, $LCI = 0$ y $LCS = D_4\overline{mr} = (3.267)(2.59) = 8.46$. La carta de control se muestra como la carta de control inferior en la figura 15-10. Esta carta de control se construyó utilizando el programa Minitab. Como ninguno de los puntos excede el límite de control superior, puede construirse ahora la carta de control para las mediciones individuales de la concentración. Si se usa un rango de movimiento de $n = 2$ observaciones, entonces $d_2 = 1.128$. Para los datos de la tabla 15-3 se tiene

$$LCS = \bar{x} + 3\frac{\overline{mr}}{d_2} = 99.1 + 3\frac{2.59}{1.128} = 105.99$$
$$LC = \bar{x} = 99.1$$
$$LCI = \bar{x} - 3\frac{\overline{mr}}{d_2} = 99.1 - 3\frac{2.59}{1.128} = 92.21$$

La carta de control para las mediciones individuales de la concentración se muestra en la carta de control superior de la figura 15-10. No hay indicios de una condición fuera de control.

Tabla 15-3 Mediciones de la concentración del proceso químico

Observación	Concentración x	Rango de movimiento mr
1	102.0	
2	94.8	7.2
3	98.3	3.5
4	98.4	0.1
5	102.0	3.6
6	98.5	3.5
7	99.0	0.5
8	97.7	1.3
9	100.0	2.3
10	98.1	1.9
11	101.3	3.2
12	98.7	2.6
13	101.1	2.4
14	98.4	2.7
15	97.0	1.4
16	96.7	0.3
17	100.3	3.6
18	101.4	1.1
19	97.2	4.2
20	101.0	3.8
	$\bar{x} = 99.1$	$\overline{mr} = 2.59$

La carta para mediciones individuales puede interpretarse en forma muy parecida a una carta de control $\bar{X}$ ordinaria. Un cambio en el promedio del proceso dará como resultado un punto (o puntos) fuera de los límites de control o bien un patrón que consiste en una serie a uno de los lados de la línea central.

Deberá tenerse cierto cuidado al interpretar patrones en la carta para rangos móviles. Los rangos móviles están correlacionados, y esta correlación con frecuencia puede inducir un patrón de corridas o ciclos en la carta. Se supone, sin embargo, que las mediciones individuales no están correlacionadas, y deberá investigarse con atención cualquier patrón aparente en la carta de control de las mediciones individuales.

La carta de control para mediciones individuales es muy sensible a los corrimientos ligeros en la media del proceso. Por ejemplo, si el tamaño del corrimiento en la media es de una desviación estándar, el número promedio de puntos para detectar este corrimiento es 43.9. Este resultado se demuestra más adelante en el capítulo. Aun cuando el desempeño de la carta de control para mediciones individuales es mucho mejor cuando los corrimientos son grandes, en muchas situaciones el corrimiento de interés no es grande y es deseable una detección mucho más rápida del mismo. En estos casos, se recomienda la *carta de control de suma acumulada* o una *carta de promedio móvil con ponderación exponencial* (Montgomery).

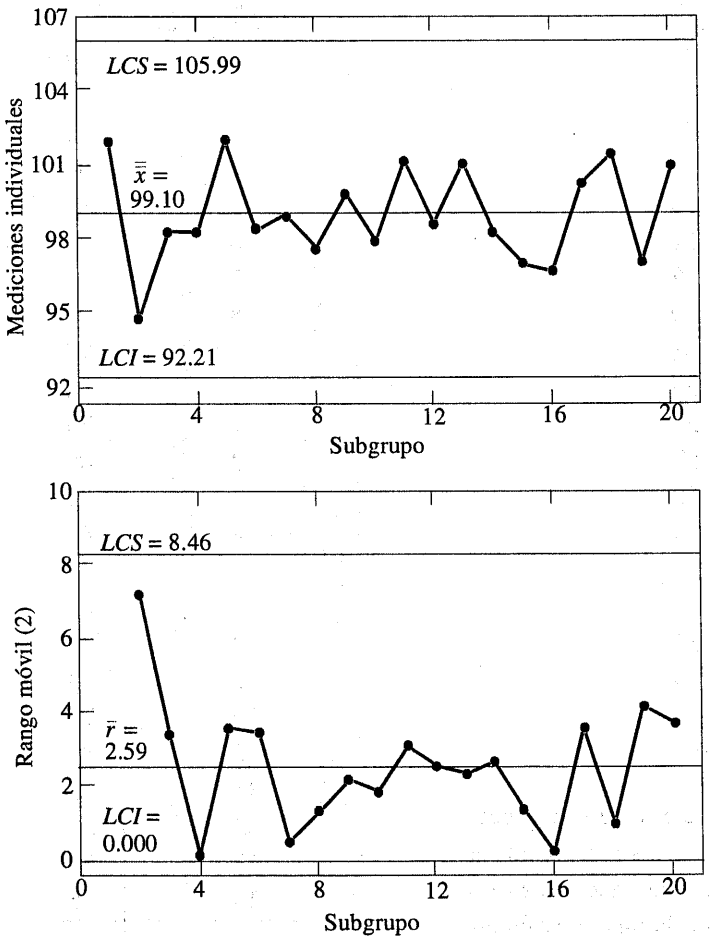


Figura 15-10 Cartas de control para mediciones individuales y para rangos móviles (del programa Minitab) para los datos de la concentración del proceso químico del ejemplo 15-2.

Algunos autores han sugerido el uso de límites con una anchura menor que 3-sigma en la carta para mediciones individuales a fin de aumentar su capacidad para detectar los corrimientos pequeños en el proceso. Esta es una sugerencia riesgosa, ya que límites más estrechos incrementarán de manera dramática las falsas alarmas de tal suerte que las cartas pueden ignorarse y volverse inútiles. Cuando haya interés en detectar corrimientos pequeños, deberán usarse las cartas de control de suma acumulada o de promedio móvil con ponderación exponencial citadas arriba.

EJERCICIOS DE LA SECCIÓN 15-6

15-8. Se hacen 20 mediciones consecutivas de la dureza de una aleación de metal, y los datos se presentan en la tabla del principio de la página siguiente.

a) Utilizando todos los datos, calcule los límites de control de prueba en las cartas para las observaciones individuales y para los rangos móviles. Construya las cartas y

Observación	Dureza
1	51
2	52
3	54
4	55
5	55
6	51
7	52
8	50
9	51
10	56
11	51
12	57
13	58
14	50
15	53
16	52
17	54
18	50
19	56
20	53

grafique los datos. Determine si el proceso está bajo control estadístico. De no ser así, suponga que pueden encontrarse las causas asignables para eliminar estas muestras, y revise los límites de control.

b) Estime la media y la desviación estándar del proceso para el proceso bajo control.

15-9. La pureza de un producto químico se mide cada hora. La determinación de la pureza de las últimas 24 horas se muestran en la siguiente tabla.

a) Utilizando todos los datos, calcule los límites de control de prueba en las cartas para las observaciones individuales y para los rangos móviles. Construya las cartas y grafique los datos. Determine si el proceso está bajo control estadístico. De no ser así, suponga que pueden encontrarse las causas asignables para eliminar estas muestras, y revise los límites de control.

b) Estime la media y la desviación estándar del proceso para el proceso bajo control.

15-10. El diámetro de unas perforaciones se mide en orden consecutivo utilizando un sensor

Observación	Pureza
1	81
2	83
3	82
4	80
5	84
6	76
7	83
8	85
9	79
10	82
11	75
12	80
13	83
14	86
15	84
16	85
17	81
18	83
19	77
20	82
21	75
22	83
23	85
24	86

automático. Los resultados de la medición de 25 perforaciones se muestran en la tabla del principio de la página siguiente.

a) Utilizando todos los datos, calcule los límites de control de prueba en las cartas para las observaciones individuales y para lcs rangos móviles. Construya la carta de control y grafique los datos. Determine si el proceso está bajo control estadístico. De no ser así, suponga que pueden encontrarse las causas asignables para eliminar estas muestras, y revise los límites de control.

b) Estime la media y la desviación estándar del proceso para el proceso bajo control.

15-11. La viscosidad de un producto químico intermedio se mide cada hora. En la segunda tabla de la página siguiente se presentan 20 muestras de tamaño $n = 1$ cada una.

a) Utilizando todos los datos, calcule los límites de control de prueba en las cartas para

Muestra	Diámetro
1	9.94
2	9.93
3	10.09
4	9.98
5	10.11
6	9.99
7	10.11
8	9.84
9	9.82
10	10.38
11	9.99
12	10.41
13	10.36
14	9.99
15	10.12
16	9.81
17	9.73
18	10.14
19	9.96
20	10.06
21	10.11
22	9.95
23	9.92
24	10.09
25	9.85

las observaciones individuales y para los rangos móviles. Construya las cartas y gra-

Muestra	Viscocidad
1	495
2	491
3	501
4	501
5	512
6	540
7	492
8	504
9	542
10	508
11	493
12	507
13	503
14	475
15	497
16	499
17	468
18	486
19	511
20	487

fique los datos. Determine si el proceso está bajo control estadístico. De no ser así, suponga que pueden encontrarse las causas asignables para eliminar estas muestras, y revise los límites de control.

b) Estime la media y la desviación estándar del proceso para el proceso bajo control.

15-7 CAPACIDAD DE UN PROCESO

Por lo general es necesario obtener información acerca de la **capacidad** del proceso; es decir, del desempeño del proceso cuando está operando bajo control. Dos herramientas gráficas, la **carta de tolerancia** (o carta de hileras) y el **histograma**, son útiles para determinar la capacidad del proceso. En la figura 15-11 se muestra la carta de tolerancia para las 20 muestras del proceso de fabricación de aspas. Las especificaciones de la abertura de las aspas son 0.5030 ± 0.0010 pulgadas. En términos de los datos codificados, el límite superior de las especificaciones es $LSE = 40$ y el límite inferior de las especificaciones es $LIE = 20$, y estos límites se muestran en la carta de la figura 15-11. Cada medición se grafica en la carta de tolerancia. Las mediciones del mismo subgrupo se unen con líneas. La carta de tolerancia es útil para revelar patrones de las mediciones individuales que ocurren con el tiempo, o puede mostrar que un valor particular de $\bar{x}$ o r fue producido por una o dos observaciones inusuales de la muestra. Por ejemplo, nótense

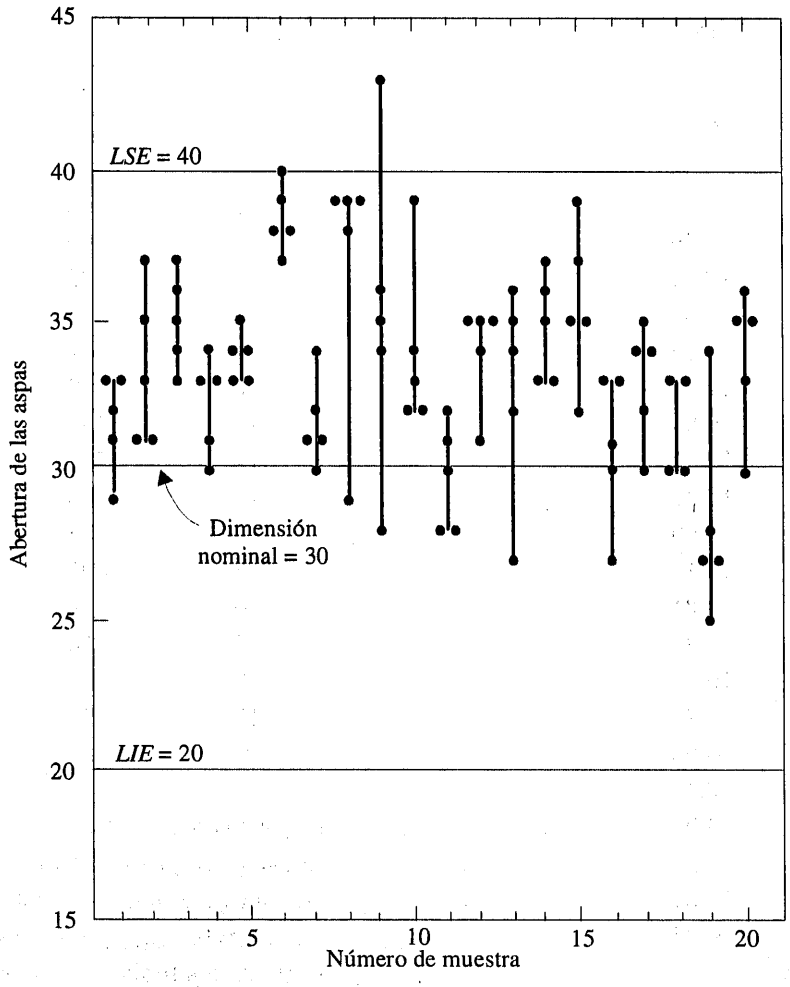


Figura 15-11 Carta de tolerancia de las aberturas de las aspas.

las dos observaciones inusuales de la muestra 9 y la observación inusual de la muestra 8. Obsérvese asimismo que es apropiado graficar los límites de las especificaciones en la carta de tolerancia, ya que se trata de una carta de mediciones individuales. **Nunca es apropiado graficar los límites de las especificaciones en una carta de control o usar las especificaciones para determinar los límites de control.** Los límites de las especificaciones y los límites de control no están relacionados. Por último, obsérvese en la figura 15-11 que el proceso está operando fuera del centro de la dimensión nominal de 30 (o 0.5030 pulgadas).

En la figura 15-12 se muestra el histograma para las mediciones de las aberturas de las aspas. Las observaciones de las muestras 6, 8, 9, 11 y 19 (que corresponden a los puntos fuera de control de las cartas $\bar{X}$ o R) se han eliminado de este histograma. La impresión general del examen de este histograma es que el proceso tiene la capacidad para cumplir con las especificaciones pero que está operando fuera del centro.

Otra forma de expresar la capacidad de un proceso es en términos del índice que se define a continuación.

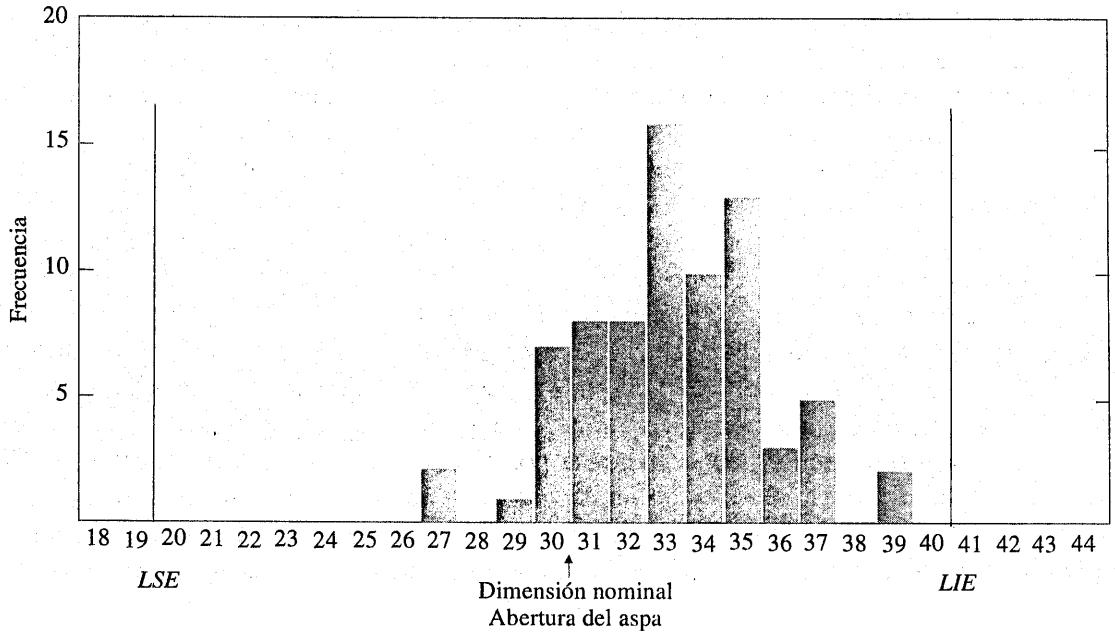


Figura 15-12 Histograma de la abertura de las aspas.

Definición

El **índice de la capacidad del proceso** (*ICP*) es

$$ICP = \frac{LSE - LIE}{6\sigma} \tag{15-10}$$

El numerador del *ICP* es la anchura de las especificaciones. A los límites 3σ a ambos lados de la media del proceso en ocasiones se les llama **límites de tolerancia natural**, ya que representan los límites que deberá satisfacer un proceso bajo control con la mayoría de las unidades producidas. Por consiguiente, es común hacer referencia a 6σ como la anchura del proceso. Para la abertura de las aspas, donde el tamaño de la muestra es 5, σ podría estimarse como

$$\hat{\sigma} = \frac{r}{d_2} = \frac{5.0}{2.326} = 2.15$$

Por lo tanto, el *ICP* se estima como

$$\begin{aligned} ICP &= \frac{LSE - LIE}{6\hat{\sigma}} \\ &= \frac{40 - 20}{6(2.15)} \\ &= 1.55 \end{aligned}$$

El *ICP* tiene una interpretación natural: (1/*ICP*)100 no es sino el porcentaje de la anchura de las especificaciones utilizado por el proceso. Por tanto, el proceso de la abertura de las aspas utiliza aproximadamente (1/1.55)100 = 64.5% de la anchura de las especificaciones.

En la figura 15-13*a* se muestra un proceso en el que el *ICP* excede la unidad. Puesto que los límites de la tolerancia natural se localizan dentro de las especificaciones, se producirán muy pocas unidades defectuosas o no conformables. Si *ICP* = 1, como se muestra en la figura 15-13*b*, resultan más unidades no conformables. De hecho, para un proceso con una distribución normal, si *ICP* = 1, la proporción de unidades no conformables es 0.27%, o 2 700 partes por millón. Por último, cuando el *ICP* es menor que la unidad, como en la figura 15-13*c*, el rendimiento del proceso es muy sensible y se producirá un gran número de unidades no conformables.

En la definición del *ICP* dada en la ecuación 15-10 se supone implícitamente que el proceso está centrado en la dimensión nominal. Si el proceso está operando fuera del centro, su **capacidad real** será menor que la indicada por el *ICP*. Es conveniente concebir el *ICP* como una

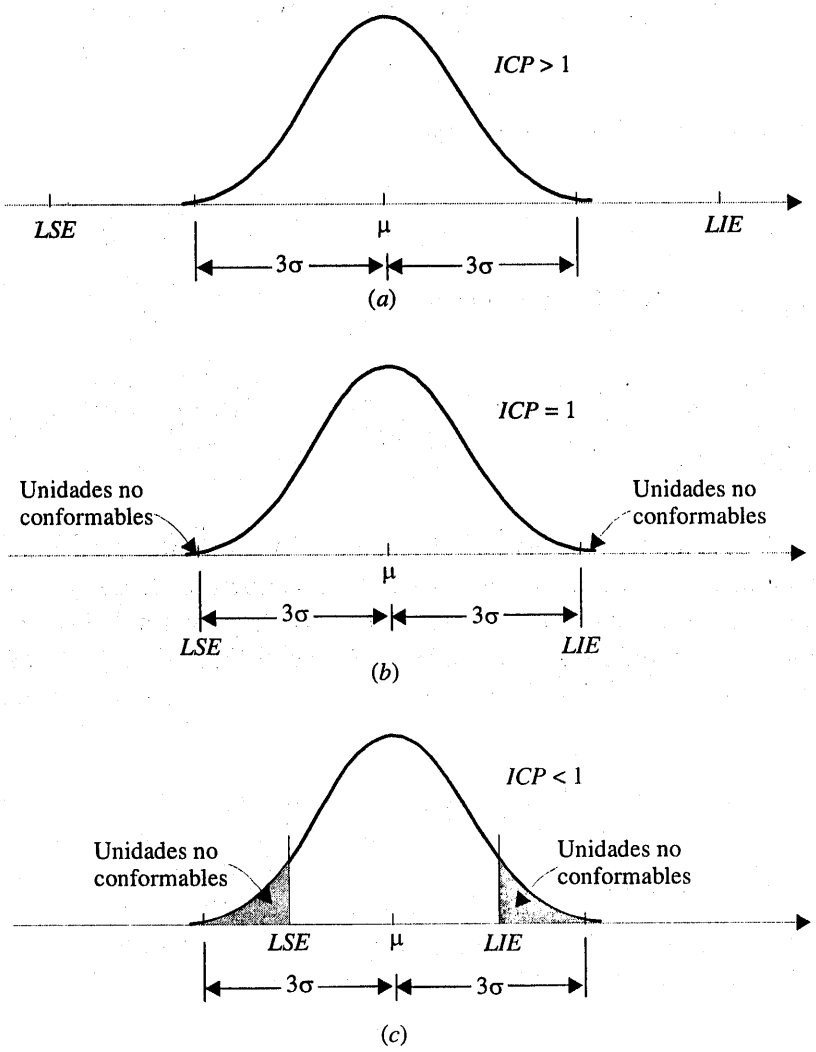


Figura 15-13 Producción incidental del proceso y el índice de la capacidad del proceso (*ICP*).

medida de la **capacidad potencial**, es decir, de la capacidad con un proceso centrado. Si el proceso no está centrado, entonces con frecuencia se usa una medida de la capacidad real. Este índice, llamado ICP_k, se define a continuación.

Definición

$$ICP_k = \text{mín}\left[\frac{LSE-\mu}{3\sigma}, \frac{\mu-LIE}{3\sigma}\right] \tag{15-11}$$

En realidad, el ICP_k es un índice de la capacidad del proceso de una cola que se calcula con relación al límite de las especificaciones más próximo a la media del proceso. Para el proceso de la abertura de las aspas, se encuentra que la estimación del índice de la capacidad del proceso ICP_k es

$$\begin{aligned}\widehat{ICP_k} &= \text{mín}\left[\frac{LSE-\bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}}-LIE}{3\hat{\sigma}}\right]\\ &= \text{mín}\left[\frac{40-33.19}{3(2.15)} = 1.06, \frac{33.19-20}{3(2.15)} = 2.04\right]\\ &= 1.06\end{aligned}$$

Obsérvese que si $ICP = ICP_k$, el proceso está centrado en la dimensión nominal. Puesto que $\widehat{ICP_k} = 1.06$ para el proceso de la abertura de las aspas e $\widehat{ICP} = 1.55$, es evidente que el proceso opera fuera del centro, como ya se observó en las figuras 15-11 y 15-12. La causa de esta operación fuera del centro se rastreó en última instancia hasta una herramienta para el moldeo de precisión. El cambio de la herramienta dio como resultado un mejoramiento sustancial del proceso (Montgomery).

Con frecuencia son motivo de interés las proporciones de la producción no conformables (o incidental) que se localizan abajo del límite inferior de las especificaciones y arriba del límite superior de las especificaciones. Suponga que la producción de un proceso con una distribución normal bajo control estadístico, se denota como X. Las fracciones se determinan a partir de

$$\begin{aligned}P(X < LIE) &= P(Z < (LIE-\mu)/\sigma)\\ P(X > LSE) &= P(Z > (LSE-\mu)/\sigma)\end{aligned}$$

EJEMPLO 15-3

En un proceso de manufactura de aparatos electrónicos una corriente tiene especificaciones de 100 ± 10 miliamperes. La media μ y la desviación estándar σ del proceso son 107.0 y 1.5, respectivamente. La media del proceso está más cerca del LSE. Por consiguiente,

$$ICP = (110-90)/(6 \cdot 1.5) = 2.22 \qquad \text{y} \qquad ICP_k = (110-107)/(3 \cdot 1.5) = 0.67$$

El valor pequeño de ICP_k indica que es probable que el proceso produzca corrientes fuera de los límites de las especificaciones. Por la distribución normal de la tabla II del apéndice

$$P(X < LIE) = P(Z < (90-107)/1.5) = P(Z < -11.33) = 0$$
$$P(X > LSE) = P(Z > (110-107)/1.5) = P(Z > 2) = 0.023$$

Para este ejemplo, la probabilidad relativamente alta de exceder el *LSE* es una advertencia de problemas potenciales con este criterio incluso si ninguna de las observaciones medidas en una muestra preliminar excede este límite. Se hace hincapié en que en el cálculo de la proporción de unidades no conformables se supone que las observaciones tienen una distribución normal y que el proceso está bajo control. Las desviaciones de la normalidad pueden afectar seriamente los resultados. El cálculo deberá interpretarse como una pauta aproximada del desempeño del proceso. Para empeorar la situación, μ y σ necesitan estimarse a partir de los datos disponibles, y un tamaño pequeño de la muestra puede resultar en estimaciones pobres que degraden aún más el cálculo.

Montgomery ofrece lineamientos para los valores apropiados del *ICP* y una tabla que relaciona la producción incidental de un proceso bajo control estadístico que tiene una distribución normal con el valor del *ICP*. Muchas compañías estadounidenses usan el valor $ICP = 1.33$ como objetivo mínimo aceptable e $ICP = 1.66$ como objetivo mínimo para la resis-tencia, la seguridad o las características críticas. Algunas compañías requieren que los procesos internos y de los proveedores alcancen un $ICP_k = 2.0$. En la figura 15-14 se ilustra un proceso con $ICP = ICP_k = 2.0$. Suponiendo una distribución normal, la producción incidental calculada para este proceso es 0.0018 partes por millón. A un proceso con $ICP_k = 2.0$ se le conoce como **proceso seis-sigma**, porque la distancia de la media del proceso a la especificación más próxima es de seis desviaciones estándar. La razón de que se requiera con frecuencia una capacidad del proceso tan alta es porque es difícil mantener la media de un proceso en el centro de las especificaciones durante periodos prolongados. Un modelo común que se usa para justificar la importancia de un proceso seis-sigma se ilustra con referencia a la figura 15-14. Si la media del proceso se corre 1.5 desviaciones estándar fuera del centro, el ICP_k disminuye a $4.5\sigma/3\sigma = 1.5$. Suponiendo un proceso con una distribución normal, la producción incidental del proceso corrido es de **3.4 partes por millón**. Por consiguiente, la media de un proceso seis-sigma puede correrse 1.5 desviaciones estándar del centro de las especificaciones y seguir manteniendo una producción incidental de 3.4 partes por millón.

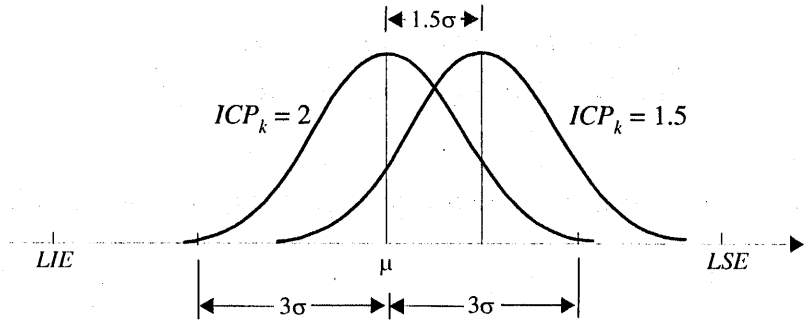


Figura 15-14 La media de un proceso seis-sigma se corre 1.5 desviaciones estándar.

Por otra parte, algunas compañías estadounidenses, en particular de la industria automotriz, han adoptado la terminología japonesa $C_p = ICP$ y $C_{pk} = ICP_k$. Debido a que C_p tiene otro significado en la estadística (en la regresión múltiple), se prefiere la notación tradicional ICP e ICP_k.

Se insiste en que los cálculos de la capacidad del proceso sólo tienen sentido para procesos estables; es decir, procesos que están bajo control. El índice de la capacidad del proceso indica si la variabilidad natural o fortuita de un proceso es aceptable o no respecto de las especificaciones.

EJERCICIOS DE LA SECCIÓN 15-7

15-12. Un proceso con una distribución normal utiliza 66.7% de la banda de las epecificaciones. Está centrado en la dimensión nominal, localizada a la mitad entre los límites superior e inferior de las especificaciones.

a) Estime el ICP y el ICP_k. Interprete estos índices.

b) ¿Qué nivel de unidades incidentales (proporción defectuosa) se produce?

15-13. Considere nuevamente el ejercicio 15-1. Use los límites de control y las estimaciones del proceso revisados.

a) Estime el ICP y el ICP_k. Interprete estos índices.

b) ¿Qué porcentaje de unidades defectuosas se está produciendo en este proceso?

15-14. Considere nuevamente el ejercicio 15-2, donde los límites de las especificaciones son 14.50 ± 0.50.

a) ¿Qué conclusiones pueden sacarse acerca de la capacidad del proceso para operar dentro de estos límites? Estime el porcentaje de artículos defectuosos que se producirán.

b) Estime el ICP y el ICP_k. Interprete estos índices.

15-15. Considere nuevamente el ejercicio 15-3. Utilizando las estimaciones del proceso, ¿cuál es el nivel de unidades incidentales si las especificaciones codificadas son 10 ± 5 mm? Estime el ICP e interprete este índice.

15-16. Un proceso con una distribución normal utiliza 85% de la banda de las especificaciones. Está centrado en la dimensión nominal, localizada a la mitad entre los límites superior e inferior de las especificaciones.

a) Estime el ICP y el ICP_k. Interprete estos índices.

b) ¿Cuál es el nivel de unidades incidentales (proporción defectuosa) que se produce?

15-17. Considere nuevamente el ejercicio 15-5. Suponga que la característica de la calidad tiene una distribución normal con las especificaciones en 220 ± 40. ¿Cuál es el nivel de unidades incidentales? Estime el ICP y el ICP_k e interprete estos índices.

15-18. Considere nuevamente el ejercicio 15-4. Suponiendo que ambas cartas presentan control estadístico y que las especificaciones del proceso están en 20 ± 5, estime ICP y ICP_k e interprete estos índices.

15-19. Considere nuevamente el ejercicio 15-7. Dado que las especificaciones están en 6.0 ± 1.0, estime el ICP y el ICP_k e interprete estos índices.

15-20. Considere nuevamente el ejercicio 15-6. ¿Cuáles son los límites de tolerancia natural de este proceso?

15-21. Considere nuevamente el ejercicio 15-11. Las especificaciones de la viscosidad están en 500 ± 25. Calcule las estimaciones de los índices de la capacidad del proceso ICP e ICP_k para este proceso y ofrezca una interpretación.

15-8 CARTAS DE CONTROL DE ATRIBUTOS

15-8.1 Carta P (carta de control de proporciones)

Muchas veces es deseable clasificar un producto como defectuoso o no defectuoso con base en la comparación con un estándar. Por lo general esta clasificación se emplea para hacer más económica y sencilla la operación de inspección. Por ejemplo, el diámetro de un rodamiento de bolas puede verificarse si se hace pasar por un escantillón que consiste en agujeros circulares perforados en una plantilla. Este tipo de medición sería mucho más simple que medir directamente el diámetro con un dispositivo tal como un micrómetro. Las cartas de control de atributos se usan en estas situaciones. Las cartas de control de atributos con frecuencia requieren tamaños de las muestras considerablemente más grandes que sus contrapartes para mediciones de variables. En esta sección se revisará la **carta de control de la proporción defectuosa**, o **carta P**. En ocasiones a la carta P se le llama **carta de control de la fracción no conformable o satisfactoria**.

Suponga que D es el número de unidades defectuosas en una muestra aleatoria de tamaño n. Se supone que D es una variable aleatoria binomial con parámetro p desconocido. La fracción defectuosa

$$\hat{P} = \frac{D}{n}$$

de cada muestra se grafica en la carta. Además, la varianza del estadístico $\hat{P}$ es

$$\sigma^2_{\hat{P}} = \frac{p(1-p)}{n}$$

Por lo tanto, una carta P de la fracción defectuosa podría construirse usando p como la línea central y los límites de control localizados en

$$LCS = p + 3\sqrt{\frac{p(1-p)}{n}}$$

$$LCI = p - 3\sqrt{\frac{p(1-p)}{n}} \tag{15-12}$$

Sin embargo, la verdadera fracción defectuosa del proceso casi siempre es desconocida, y debe estimarse utilizando los datos de muestras preliminares.

Suponga que se cuenta con m muestras preliminares de tamaño n cada una, y sea D_i el número de unidades defectuosas en la muestra i-ésima. $\hat{P}_i = D_i/n$ es la proporción muestral de unidades defectuosas en la muestra i-ésima. La fracción defectuosa promedio es

$$\overline{P} = \frac{1}{m}\sum_{i=1}^{m}\hat{P}_i = \frac{1}{mn}\sum_{i=1}^{m}D_i \tag{15-13}$$

Entonces $\overline{P}$ puede usarse como estimador de p en la línea central y en los cálculos de los límites de control.

Definición

La línea central y los límites de control superior e inferior de la carta P son

$$LCS = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$LC = \bar{p}$$

$$LCI = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \qquad (15\text{-}14)$$

donde $\bar{p}$ es el valor observado de la fracción defectuosa promedio.

Estos límites de control se basan en la aproximación normal de la distribución binomial. Cuando el valor de p es pequeño, la aproximación normal quizá no siempre sea adecuada. En tales casos, pueden usarse los límites de control obtenidos directamente de una tabla de probabilidades binomiales. Cuando $\bar{p}$ es pequeño, el límite de control inferior obtenido de la aproximación normal puede ser un número negativo. Si esto ocurre, se acostumbra considerar cero el límite de control inferior.

EJEMPLO 15-4

Suponga que quiere construirse una carta de control de la fracción defectuosa para una línea de producción de sustrato cerámico. Se tienen 20 muestras preliminares, cada una de tamaño 100; el número de unidades defectuosas de cada muestra se presenta en la tabla 15-4. Suponga que las muestras están numeradas en la secuencia de producción. Obsérvese que $\bar{p}$ = (800/2 000) = 0.40; por lo tanto, los parámetros de prueba para la carta de control son

$$LCS = 0.40 + 3\sqrt{\frac{(0.40)(0.60)}{100}} = 0.55$$

$$LC = 0.40$$

$$LCI = 0.40 - 3\sqrt{\frac{(0.40)(0.60)}{100}} = 0.25$$

La carta de control se presenta en la figura 15-15. Todas las muestras están bajo control. En caso de que no lo estuvieran, se buscarían las causas asignables de la variación y se haría la revisión correspondiente de los límites. Esta carta puede usarse para controlar la producción futura.

Aun cuando este proceso presenta control estadístico, su proporción de unidades defectuosas ($\bar{p}$ = 0.40) es muy pobre. Deberán tomarse las medidas necesarias para investigar el proceso y determinar por qué se está produciendo un número tan alto de unidades defectuosas. Las

Tabla 15-4 Número de unidades defectuosas en muestras de 100 sustratos cerámicos

Muestra	Número de unidades defectuosas	Muestra	Número de unidades defectuosas
1	44	11	36
2	48	12	52
3	32	13	35
4	50	14	41
5	29	15	42
6	31	16	30
7	46	17	46
8	52	18	38
9	44	19	26
10	48	20	30

unidades defectuosas deberán analizarse para determinar los tipos de defectos específicos que presentan. Una vez que se conozcan los tipos de defectos, deberán investigarse los cambios en el proceso para determinar su impacto en los niveles de los defectos. Los experimentos diseñados pueden ser útiles a este respecto.

El software de computadora también produce una **carta *NP***. Se trata de una carta de control de $n\hat{P} = D$, el número de unidades defectuosas en una muestra. Los puntos, la línea central y los límites de control de esta carta son múltiplos (por n) de los elementos correspondientes de una carta P. El uso de una carta NP evita las fracciones de una carta P.

15-8.2 Carta U (carta de control de defectos por unidad)

En ocasiones es necesario monitorear el número de defectos en una unidad de un producto en lugar de la proporción defectuosa. Suponga que en la producción de una tela es necesario controlar el número de defectos por yarda, o que en el ensamblaje del ala de un avión debe contro-

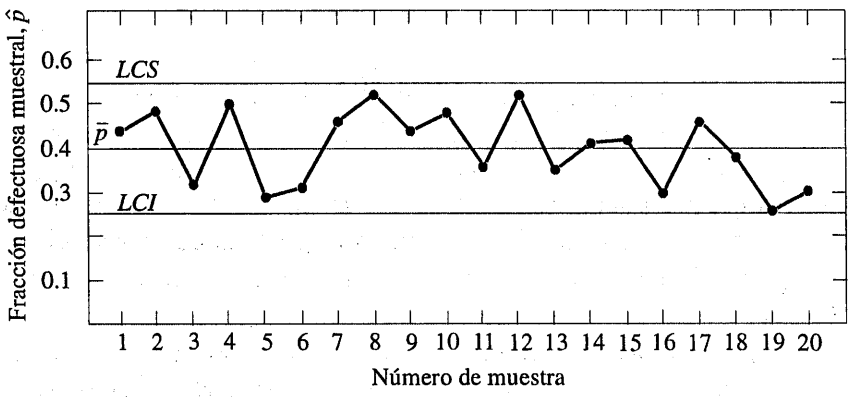


Figura 15-15 Carta P para un sustrato cerámico.

larse el número de remaches faltantes. En estas situaciones puede usarse la carta de control de defectos por unidad o **carta *U***. Muchas situaciones de defectos por unidad pueden modelarse con la distribución de Poisson.

Si cada muestra consta de n unidades y hay C defectos en total en la muestra, entonces

$$U = \frac{C}{n}$$

es el número promedio de defectos por unidad. Puede construirse una carta U para estos datos. Si hay m muestras preliminares, y el número de defectos por unidad en estas muestras son U_1, U_2, ..., U_m, entonces el estimador del número promedio de defectos por unidad es

$$\bar{U} = \frac{1}{m}\sum_{i=1}^{m} U_i \tag{15-15}$$

Los parámetros de la carta U se definen como sigue.

Definición

La línea central y los límites de control superior e inferior de la carta U son

$$\begin{aligned} LCS &= \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} \\ LC &= \bar{u} \\ LCI &= \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} \end{aligned} \tag{15-16}$$

donde $\bar{u}$ es el número promedio de defectos por unidad.

Si el número de defectos en una unidad es una variable aleatoria de Poisson con parámetro λ, entonces la media y la varianza de esta distribución son ambas λ. Cada punto de la carta es U, el número promedio de defectos por unidad de la muestra de n unidades. Por lo tanto, la media de U es λ y la varianza de U es λ/n. Por lo general no se conoce λ, y $\bar{U}$ es el estimador de λ que se usa para establecer los límites de control.

Estos límites de control se basan en la aproximación normal de la distribución de Poisson. Cuando λ es pequeña, la aproximación normal quizá no sea siempre adecuada. En tales casos, pueden usarse los límites de control obtenidos directamente de una tabla de probabilidades de Poisson. Si $\bar{u}$ es pequeño, el límite de control inferior obtenido de la aproximación normal puede ser un número negativo. Si esto ocurre, se acostumbra considerar cero el límite de control inferior.

EJEMPLO 15-5

Se ensamblan tarjetas de circuitos impresos mediante una combinación de ensamblaje manual y automatizado. Se usa una máquina de soldadura líquida para hacer las conexiones mecánicas y eléctricas de los componentes integrados en la tarjeta. Las tarjetas pasan por el proceso de soldadura líquida de manera casi continua, y cada hora se seleccionan e inspeccionan cinco tarjetas para fines de control del proceso. Se registra el número de defectos en cada muestra de cinco tarjetas. En la tabla 15-5 se muestran los resultados de 20 muestras.

La línea central de la carta U es

$$\bar{u} = \frac{1}{20}\sum_{i=1}^{20} u_i = \frac{32}{20} = 1.6$$

y los límites de control superior e inferior son

$$LCS = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 1.6 + 3\sqrt{\frac{1.6}{5}} = 3.3$$

$$LCI = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 1.6 - 3\sqrt{\frac{1.6}{5}} < 0$$

La carta de control se grafica en la figura 15-16. Como LCI es negativo, se hace igual a 0. En la carta de control de la figura 15-16 se observa que el proceso está bajo control. Sin embargo, ocho defectos en cada grupo de cinco tarjetas de circuitos son demasiados (aproximadamente 8/5 = 1.6 defectos/tarjeta), y se requiere mejorar el proceso. Es necesario hacer una investigación de los tipos específicos de defectos encontrados en las tarjetas de circuitos impresos. Esto por lo general sugiere caminos potenciales para el mejoramiento del proceso.

Tabla 15-5 Número de defectos en muestras de cinco tarjetas de circuitos impresos

Muestra	Número de defectos	Defectos por unidad u_i	Muestra	Número de defectos	Defectos por unidad u_i
1	6	1.2	11	9	1.8
2	4	0.8	12	15	3.0
3	8	1.6	13	8	1.6
4	10	2.0	14	10	2.0
5	9	1.8	14	8	1.6
6	12	2.4	16	2	0.4
7	16	3.2	17	7	1.4
8	2	0.4	18	1	0.2
9	3	0.6	19	7	1.4
10	10	2.0	20	13	2.6

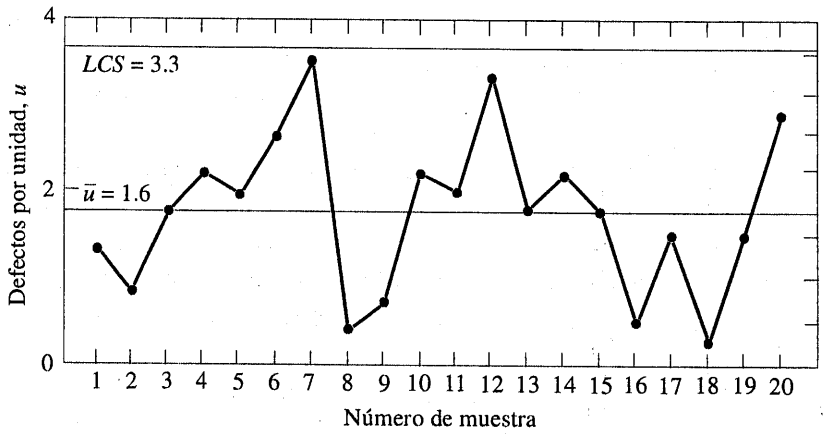


Figura 15-16 Carta U de los defectos por unidad en las tarjetas de circuitos impresos.

El software de computadora también produce una **carta C**. Se trata de una carta de control de C, el total de defectos en una muestra. Los puntos, la línea central y los límites de control para esta carta son múltiplos (por n) de los elementos correspondientes de una carta U. El uso de una carta C evita las fracciones que pueden ocurrir en una carta U.

EJERCICIOS DE LA SECCIÓN 15-8

15-22. Suponga que se han encontrado las siguientes fracciones defectuosas en muestras consecutivas de tamaño 100 (léase hacia abajo):

0.09	0.03	0.12
0.10	0.05	0.14
0.13	0.13	0.06
0.08	0.10	0.05
0.14	0.14	0.14
0.09	0.07	0.11
0.10	0.06	0.09
0.15	0.09	0.13
0.13	0.08	0.12
0.06	0.11	0.09

a) Utilizando todos los datos, calcule los límites de control de prueba para una carta de control de la fracción defectuosa, construya la carta y grafique los datos.

b) Determine si el proceso está bajo control estadístico. De no ser así, suponga que pueden encontrarse las causas asignables y que los puntos fuera de control pueden eliminarse. Revise los límites de control.

15-23. Los datos siguientes representan el número de defectos de soldadura observados en 24 muestras de cinco tarjetas de circuitos impresos: 7, 6, 8, 10, 24, 6, 5, 4, 8, 11, 15, 8, 4, 16, 11, 12, 8, 6, 5, 9, 7, 14, 8, 21.

a) Utilizando todos los datos, calcule los límites de control de prueba para una carta de control U, construya la carta y grafique los datos.

b) ¿Puede concluirse que el proceso está bajo control usando una carta U? De no ser así, suponga que pueden encontrarse las causas asignables, enliste los puntos y revise los límites de control.

15-24. Los datos siguientes representan el número de defectos por 1 000 pies de alambre con forro de hule: 1, 1, 3, 7, 8, 10, 5, 13, 0, 19, 24, 6, 9, 11, 15, 8, 3, 6, 7, 4, 9, 20, 11, 7, 18, 10, 6, 4, 0, 9, 7, 3, 1, 8, 12. ¿Los datos provienen de un proceso bajo control?

15-25. Considere los datos del ejercicio 15-23. Establezca una carta C para este proceso. Compárela con la carta U del ejercicio 15-23. Comente los resultados.

15-26. Los datos siguientes corresponden al número de juntas de soldadura defectuosas encontradas en muestras consecutivas de 500 juntas de soldadura.

Día	No. de juntas defectuosas	Día	No. de juntas defectuosas
1	106	12	37
2	116	13	25
3	164	14	88
4	89	15	101
5	99	16	64
6	40	17	51
7	112	18	74
8	36	19	71
9	69	20	43
10	74	21	80
11	42		

a) Utilizando todos los datos, calcule los límites de control de prueba de una carta de control para fracciones defectuosas, construya la carta y grafique los datos.

b) Determine si el proceso está bajo control estadístico. De no ser así, suponga que pueden encontrarse las causas asignables y que los puntos fuera de control pueden eliminarse. Revise los límites de control.

15-9 DESEMPEÑO DE LAS CARTAS DE CONTROL

Especificar los límites de control es una de las decisiones críticas que deben tomarse cuando se diseña una carta de control. Al alejar los límites de control de la línea central, se reduce el riesgo de un error tipo I; es decir, el riesgo de que un punto se localice fuera de los límites de control, indicando una condición fuera de control, cuando no hay ninguna causa asignable presente. Sin embargo, al ensanchar los límites de control también se incrementará el riesgo del error tipo II; es decir, el riesgo de un punto que se localice entre los límites de control cuando el proceso en realidad se encuentra fuera de control. Si los límites de control se acercan a la línea central, se obtiene el efecto contrario: el riesgo del error tipo I se incrementa, mientras que el riesgo del error tipo II se reduce.

Los límites de control en una carta de control de Shewhart se localizan generalmente a una distancia de más o menos tres desviaciones estándar desde la línea central de la variable graficada en la carta. Es decir, la constante k de la ecuación 15-1 deberá ser igual a 3. Estos límites se llaman **límites de control 3-sigma**.

Una forma de evaluar las decisiones respecto del tamaño de la muestra y la frecuencia de muestreo es a través de la **longitud de la corrida promedio** (**ARL**, por sus siglas en inglés) de la carta de control. En esencia, la ARL es el número promedio de puntos que deben graficarse antes de que un punto indique una condición fuera de control. Para cualquier carta de control de Shewhart, la ARL puede calcularse a partir de la media de una variable aleatoria geométrica (Montgomery). Suponga que p es la probabilidad de que cualquier punto exceda los límites de control. Entonces

$$ARL = \frac{1}{p} \tag{15-17}$$

Por tanto, para una carta $\overline{X}$ con límites 3-sigma, $p = 0.0027$ es la probabilidad de que un solo punto se localice fuera de los límites cuando el proceso está bajo control, de donde

$$ARL = \frac{1}{p} = \frac{1}{0.0027} \cong 370$$

es la longitud de la corrida promedio de la carta $\overline{X}$ cuando el proceso está bajo control. Es decir, incluso si el proceso se mantiene bajo control, una señal de fuera de control se generará cada 370 puntos, en promedio.

Considérese el proceso de los anillos para pistones analizado antes, y suponga que se hace un muestreo cada hora. Por tanto, se tendrá una **falsa alarma** aproximadamente cada 370 horas en el promedio. Suponer que se está usando un tamaño de la muestra de $n = 5$ y que cuando el proceso se sale de control la media se corre a 74.0135 mm. Entonces, la probabilidad de que $\overline{X}$ se localice entre los límites de control de la figura 15-3 es igual a

$$\begin{aligned} &P[73.9865 \le \overline{X} \le \text{cuando } \mu = 74.0135] \\ &= P\left[\frac{73.9865 - 74.0135}{0.0045} \le Z \le \frac{74.0135 - 74.0135}{0.0045}\right] \\ &= P[-6 \le Z \le 0] = 0.5 \end{aligned}$$

Por lo tanto, p en la ecuación 15-17 es 0.50, y la ARL de fuera de control es

$$ARL = \frac{1}{p} = \frac{1}{0.5} = 2$$

Es decir, la carta de control requerirá, en promedio, dos muestras para detectar el corrimiento en el proceso, por lo que transcurrirán dos horas entre el corrimiento y su detección (*de nueva cuenta, en el promedio*). Suponga que este enfoque sea inaceptable, debido a que la producción de anillos para pistones con un diámetro promedio de 74.0135 mm resulta en costos excesivos de unidades de desecho y demoras en el ensamblaje final de los motores. ¿Cómo puede reducirse el tiempo necesario para detectar la condición fuera de control? Un método es tomar muestras más seguido. Por ejemplo, si se toman muestras cada media hora, entonces sólo transcurrirá una hora (en promedio) entre el corrimiento y su detección. La segunda posibilidad es aumentar el tamaño de la muestra. Por ejemplo, si se usa $n = 10$, entonces los límites de control de la figura 15-3 se estrecharán a 73.9905 y 74.0095. La probabilidad de que $\overline{X}$ se localice entre los límites de control cuando la media del proceso es 74.0135 mm es aproximadamente 0.1, de donde $p = 0.9$, y la ARL de fuera de control es

$$ARL = \frac{1}{p} = \frac{1}{0.9} = 1.11$$

Tabla 15-6 Longitud de la serie promedio (ARL) para una carta $\bar{X}$ con límites de control 3-sigma

Magnitud del cambio en el proceso	ARL $n = 1$	ARL $n = 4$
0	370.4	370.4
0.5σ	155.2	43.9
1.0σ	43.9	6.3
1.5σ	15.0	2.0
2.0σ	6.3	1.2
3.0σ	2.0	1.0

Por tanto, el tamaño más grande de la muestra permitiría que el corrimiento se detectara casi con el doble de la rapidez que con el tamaño anterior. Si llegara a ser importante detectar el corrimiento en la primera hora después de que ocurra, dos diseños de cartas de control funcionarían:

Diseño 1	Diseño 2
Tamaño de la muestra: $n = 5$	Tamaño de la muestra: $n = 10$
Frecuencia de muestreo: cada media hora	Frecuencia de muestreo: cada hora

En la tabla 15-6 se presentan las longitudes de la corrida promedio para una carta $\bar{X}$ con límites de control 3-sigma. Las longitudes de la corrida promedio se calculan para corrimientos en la media del proceso de 0 a 3.0σ y para tamaños de la muestra de $n = 1$ y $n = 4$ utilizando $1/p$, donde p es la probabilidad de que un punto se localice fuera de los límites de control. En la figura 15-17 se ilustra un corrimiento en la media del proceso de 2σ.

EJERCICIOS DE LA SECCIÓN 15-9

15-27. Considérese la carta de control $\bar{X}$ de la figura 15-3. Suponga que la media se corre a 74.010 mm.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-28. En una carta $\bar{X}$ se utilizan muestras de tamaño 4. La línea central está en 100, y los límites de control 3-sigma superior e inferior están en 106 y 94, respectivamente.

a) ¿Cual es σ del proceso?

b) Suponga que la media del proceso se corre a 105. Encuentre la probabilidad de

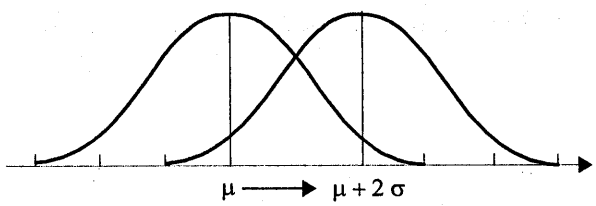


Figura 15-17 Corrimiento en la medida del proceso de 2σ.

que este corrimiento se detectará en la siguiente muestra.

c) Encuentre la ARL para detectar el corrimiento del inciso b).

15-29. Considere la carta de control $\bar{X}$ revisada del ejercicio 15-1 con $\hat{\sigma} = 2.466$, $LCS = 37.404$, $LCI = 30.780$ y $n = 5$. Suponga que la media se corre a 38.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-30. Considérese la carta de control $\bar{X}$ del ejercicio 15-2 con $\bar{r} = 0.344$, $LCS = 14.708$, $LCI = 14.312$ y $n = 5$. Suponga que la media se corre a 14.6.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-31. Considérese la carta de control $\bar{X}$ del ejercicio 15-3 con $\bar{r} = 6.750$, $LCS = 15.630$, $LCI = 5.795$ y $n = 4$. Suponga que la media se corre a 14.5.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-32. Considérese la carta de control $\bar{X}$ del ejercicio 15-4 con $\hat{\sigma} = 1.40$, $LCS = 21.88$, $LCI = 18.12$ y $n = 5$. Suponga que la media se corre a 17.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-33. Considérese la carta de control $\bar{X}$ del ejercicio 15-5 con $\bar{r} = 34.286$, $LCS = 242.780$, $LCI = 203.220$ y $n = 5$. Suponga que la media se corre a 215.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-34. Considérese la carta de control $\bar{X}$ revisada del ejercicio 15-6 con $\hat{\sigma} = 0.000924$, $LCS = 0.0635$, $LCI = 0.624$ y $n = 5$. Suponga que la media se corre a 0.0625.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-35. Considérese la carta de control $\bar{X}$ revisada del ejercicio 15-7 con $\hat{\sigma} = 0.669$, $LCS = 7.443$, $LCI = 5.125$ y $n = 3$. Suponga que la media se corre a 5.5.

a) ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra?

b) ¿Cuál es la ARL después del corrimiento?

15-10 CARTA DE CONTROL DE SUMA ACUMULADA

En las secciones 15-5 y 15-6 se han presentado los tipos básicos de las cartas de control de Shewhart. Una de las principales desventajas de cualquier carta de control de Shewhart es que la carta es relativamente insensible a cambios pequeños en el proceso, por ejemplo, del orden de aproximadamente 1.5σ o menos. Una razón de este desempeño relativamente pobre para detectar cambios pequeños en el proceso es que la carta de Shewhart tan sólo hace uso de la información del último punto graficado, e ignora la información de la secuencia de puntos. Este problema puede resolverse, hasta cierto punto, agregando criterios tales como las reglas Western Electric a una carta de Shewhart, pero el uso de estas reglas reduce la simplicidad y facilidad de interpretación de la carta. Estas reglas también ocasionarían que la longitud de la corrida promedio bajo control de una carta de Shewhart quede por debajo de 370. Este incremento del índice de falsas alarmas puede tener consecuencias serias en la práctica.

Una alternativa muy efectiva para la carta de control de Shewhart es la carta de control de **suma acumulada** (o **CUSUM**, por sus siglas en inglés). Esta carta tiene un desempeño mucho mejor que la carta de Shewhart (en términos de la ARL) para detectar cambios pequeños, pero no produce una reducción significativa de la ARL bajo control. En esta sección se ilustrará el uso de la suma acumulada (CUSUM) para promedios muestrales y mediciones individuales.

En la carta CUSUM se grafica las suma acumulada de las desviaciones de los valores muestrales del valor objetivo. Por ejemplo, suponer que se colectan muestras de tamaño $n \geq 1$, y que $\bar{X}_j$ es el promedio de la muestra j-ésima. Entonces si μ_0 es el objetivo para la media del proceso, la carta de control de suma acumulada CUSUM se forma graficando la cantidad

$$S_i = \sum_{j=1}^{i} (\bar{X}_j - \mu_0) \tag{15-18}$$

contra el número de muestra i. Ahora bien, a S_i se le llama la suma acumulada hasta e incluyendo la muestra i-ésima. Debido a que se combina información de *varias* muestras, las cartas de suma acumulada son más eficaces que las cartas de Shewhart para detectar corrimientos pequeños en el proceso. Además, son particularmente eficaces con muestras de tamaño $n = 1$. Esto hace que la carta de control de suma acumulada sea un buen candidato para usarla en las industrias química y de procesamiento donde los subgrupos racionales muchas veces son de tamaño 1, así como en la fabricación de piezas discretas con medición automática de cada pieza y control en línea utilizando directamente una microcomputadora en el centro de trabajo.

Si el proceso permanece bajo control en el valor objetivo μ_0, la suma acumulada definida en la ecuación 15-18 deberá fluctuar alrededor de cero. Sin embargo, si la media se corre hacia arriba a algún valor $\mu_1 > \mu_0$, por ejemplo, entonces se desarrollará una tendencia ascendente o positiva en la suma acumulada S_i. Recíprocamente, si la media cambia hacia abajo a algún valor $\mu_1 < \mu_0$, entonces se desarrollará una tendencia descendente o negativa en S_i. Por lo tanto, si se desarrolla una tendencia ascendente o descendente en los puntos graficados, esto deberá considerarse como evidencia de que la media del proceso se ha corrido, y deberá llevarse a cabo la búsqueda de una causa asignable.

Esta teoría puede demostrarse fácilmente aplicando la CUSUM a los datos de la concentración del proceso químico del ejemplo 15-2, tabla 15-4. Puesto que las lecturas de la concentración son mediciones individuales, se tomaría $\bar{X}_j = X_j$ al calcular la CUSUM. Suponga que el valor objetivo para la concentración es $\mu_0 = 99$. Entonces la CUSUM de la ecuación 15-18 es

$$\begin{aligned} S_i &= \sum_{j=1}^{i} (X_j - 99) \\ &= (X_i - 99) + \sum_{j=1}^{i-1} (X_j - 99) \\ &= (X_i - 99) + S_{i-1} \end{aligned}$$

En la tabla 15-7 se muestra el cálculo de esta CUSUM, donde el valor inicial de la CUSUM, S_0, se toma como cero. En la figura 15-18 se grafica la CUSUM de la última columna de la tabla 15-7. Obsérvese que la CUSUM fluctúa alrededor del valor 0.

Tabla 15-7 Cálculos de la CUSUM para los datos de la concentración del proceso químico de la tabla 15-3

Observación, i	x_i	$x_i - 99$	$s_i = (x_i - 99) + s_{i-1}$
1	102.0	3.0	3.0
2	94.8	–4.2	–1.2
3	98.3	–0.7	–1.9
4	98.4	–0.6	–2.5
5	102.0	3.0	0.5
6	98.5	–0.5	0.0
7	99.0	0.0	0.0
8	97.7	–1.3	–1.3
9	100.0	1.0	–0.3
10	98.1	–0.9	–1.2
11	101.3	2.3	1.1
12	98.7	–0.3	0.8
13	101.1	2.1	2.9
14	98.4	–0.6	2.3
15	97.0	–2.0	0.3
16	96.7	–2.3	–2.0
17	100.3	1.3	–0.7
18	101.4	2.4	1.7
19	97.2	–1.8	–0.1
20	101.0	2.0	1.9

La gráfica de la figura 15-18 no es una carta de control porque no tiene límites de control. Hay dos enfoques generales para determinar los límites de control de CUSUM. El más antiguo de estos dos métodos es el procedimiento de la máscara V. En la figura 15-19*a* se ilustra una máscara V típica. Se trata de una muesca en forma de "V" en un plano que puede colocarse en diferentes lugares de la carta de CUSUM. El procedimiento de decisión consiste en colocar la máscara V sobre la carta de control de suma acumulada con el punto O sobre el último valor de s_i y la línea OP paralela al eje horizontal. Si todas las sumas acumuladas previas, $s_1, s_2, \ldots, s_{i-1}$ se localizan dentro de los dos brazos de la máscara V, el proceso está bajo control. Sin embargo, si cualquiera de las s_i se localiza fuera de los brazos de la máscara, el proceso se considera fuera de control. En una aplicación real, la máscara V se aplicaría a cada nuevo punto de la carta CUSUM tan pronto como se haya graficado. En el ejemplo presentado en la figura 15-19*b*, aparece indicado un corrimiento ascendente en la media, ya que al menos uno de los puntos que han ocurrido antes de la muestra 22 se localiza ahora abajo del brazo inferior de la máscara, cuando la máscara V se centra en la observación número 30. Si el punto se localiza arriba del brazo superior, esto indicaría un corrimiento descendente en la media. Por tanto, la máscara V forma un marco de referencia visual similar a los límites de control en una carta de control de Shewhart ordinaria. Para los detalles técnicos del diseño de una máscara V, ver Montgomery.

Aun cuando algunos programas de computadora grafican las sumas acumuladas con el esquema de control de la máscara V, los autores sienten que el otro enfoque del control de la

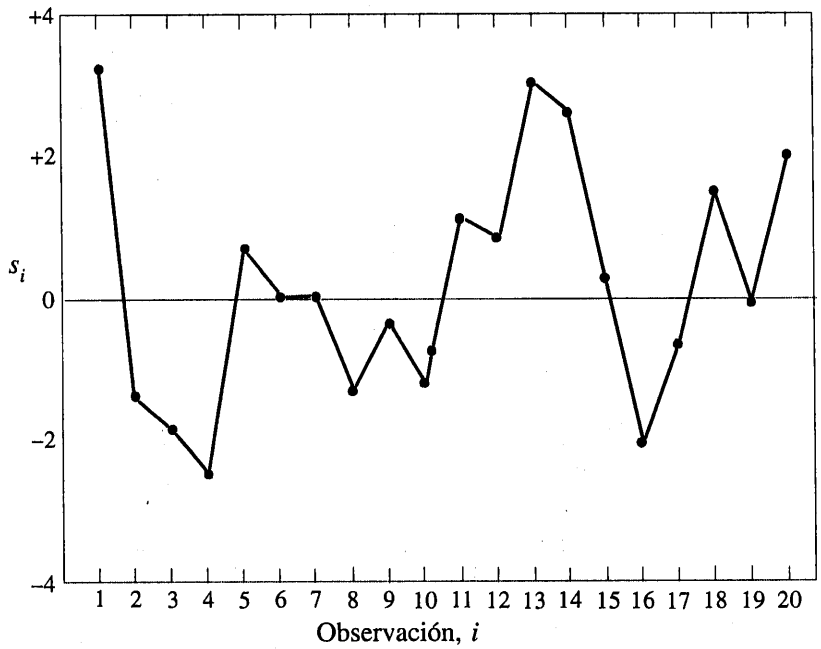


Figura 15-18 Gráfica de la suma acumulada para los datos de la concentración, tabla 15-7.

suma acumulada, la **CUSUM tabular**, es superior. El procedimiento tabular es particularmente atractivo cuando la CUSUM se implementa en una computadora.

Sea $S_H(i)$ una CUSUM de cola superior para el periodo i y sea $S_L(i)$ una CUSUM de cola inferior para el periodo i. Estas cantidades se calculan a partir de

$$s_H(i) = \text{máx}[0, \bar{x}_i - (\mu_0 + K) + s_H(i-1)] \tag{15-19}$$

y

$$s_L(i) = \text{máx}[0, (\mu_0 - K) - \bar{x}_i + s_L(i-1)] \tag{15-20}$$

respectivamente, donde los valores iniciales $s_H(0) = s_L(0) = 0$. En las ecuaciones 15-19 y 15-20 a K se le llama el **valor de referencia**, el cual por lo general se escoge cerca de la mitad entre el objetivo μ_0 y el valor de la media que corresponde al estado fuera de control, $\mu_1 = \mu_0 + \Delta$. Es decir, K es aproximadamente la mitad de la magnitud del cambio en el que se tiene interés, o

$$K = \frac{\Delta}{2}$$

Obsérvese que $S_H(i)$ y $S_L(i)$ acumulan las desviaciones del valor objetivo que son mayores que K, con ambas cantidades reinicializadas en cero cuando asumen valores negativos. Si $S_H(i)$ o $S_L(i)$ exceden una constante H, el proceso está fuera de control. A esta constante H suele llamársele el **intervalo de decisión**.

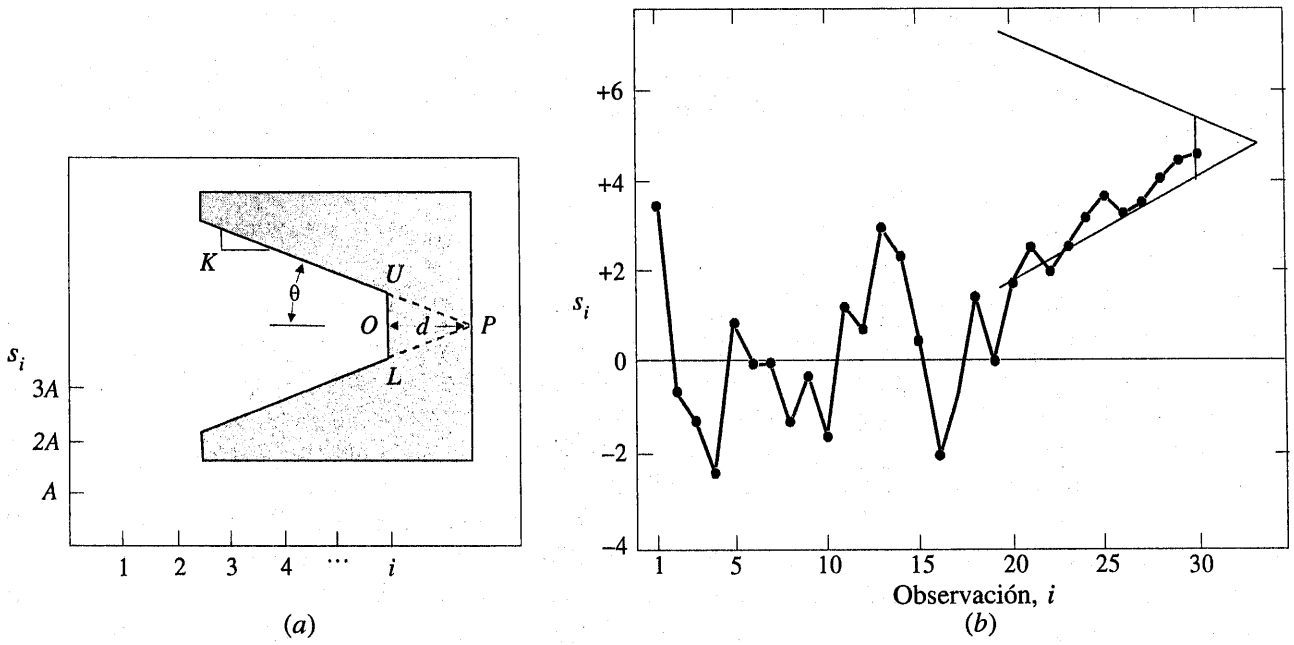


Figura 15-19 La carta de control de suma acumulada. ***a***) La máscara V y escalación. ***b***) La carta de control de suma acumulada en operación.

EJEMPLO 15-6

Una CUSUM tabular

Se ilustrará la CUSUM tabular aplicándola a los datos de la concentración del proceso químico de la tabla 15-7. El objetivo del proceso es $\mu_0 = 99$, y se usará $K = 1$ como valor de referencia y $H = 10$ como intervalo de decisión. Las razones de estas elecciones se explicarán más adelante.

En la tabla 15-8 se muestra el esquema de la CUSUM tabular para los datos de la concentración del proceso químico. Para ilustrar los cálculos, obsérvese que

$$\begin{aligned} s_H(i) &= \text{máx}[0, x_i - (\mu_0 + K) + s_H(i-1)] \\ &= \text{máx}[0, x_i - (99+1) + s_H(i-1)] \\ &= \text{máx}[0, x_i - 100 + s_H(i-1)] \end{aligned}$$

$$\begin{aligned} s_L(i) &= \text{máx}[0, (\mu_0 - K) - x_i + s_L(i-1)] \\ &= \text{máx}[0, (99-1) - x_i + s_L(i-1)] \\ &= \text{máx}[0, 98 - x_i + s_L(i-1)] \end{aligned}$$

Por lo tanto, para la observación 1 las CUSUM son

$$\begin{aligned} s_H(1) &= \text{máx}[0, x_1 - 100 + s_H(0)] \\ &= \text{máx}[0, 102.0 - 100 + 0] \\ &= 2.0 \end{aligned}$$

Tabla 15-8 La CUSUM tabular para los datos de la concentración del proceso químico

Observación	CUSUM superior				CUSUM inferior		
i	x_i	$x_i - 100$	$s_H(i)$	n_H	$98 - x_i$	$s_L(i)$	n_L
1	102.0	2.0	2.0	1	–4.0	0.0	0
2	94.8	–5.2	0.0	0	3.2	3.2	1
3	98.3	–1.7	0.0	0	–0.3	2.9	2
4	98.4	–1.6	0.0	0	–0.4	2.5	3
5	102.0	2.0	2.0	1	–4.0	0.0	0
6	98.5	–1.5	0.5	2	–0.5	0.0	0
7	99.0	–1.0	0.0	0	–1.0	0.0	0
8	97.7	–2.3	0.0	0	0.3	0.3	1
9	100.0	0.0	0.0	0	–2.0	0.0	0
10	98.1	–1.9	0.0	0	–0.1	0.0	0
11	101.3	1.3	1.3	1	–3.3	0.0	0
12	98.7	–1.3	0.0	0	–0.7	0.0	0
13	101.1	1.1	1.1	1	–3.1	0.0	0
14	98.4	–1.6	0.0	0	–0.4	0.0	0
15	97.0	–3.0	0.0	0	1.0	1.0	1
16	96.7	–3.3	0.0	0	1.3	2.3	2
17	100.3	0.3	0.3	1	–2.3	0.0	0
18	101.4	1.4	1.7	2	–3.4	0.0	0
19	97.2	–2.8	0.0	0	0.8	0.8	1
20	101.0	1.0	1.0	0	–3.0	0.0	0

y

$$\begin{aligned} s_L(1) &= \text{máx}[0, 98 - x_1 + s_L(0)] \\ &= \text{máx}[0, 98 - 102.0 + 0] \\ &= 0 \end{aligned}$$

como se muestra en la tabla 15-8. Las cantidades n_H y n_L de la tabla 15-8 indican el número de periodos que las CUSUM $s_H(i)$ o $s_L(i)$ han sido diferentes de cero. Obsérvese que las CUSUM de este ejemplo nunca exceden el intervalo de decisión $H = 10$. Se concluiría por lo tanto que el proceso está bajo control.

Cuando la CUSUM tabular indica que el proceso está fuera de control, deberá buscarse la causa asignable, adoptar las acciones correctivas pertinentes y reinicializar las CUSUM en cero. Puede ser útil tener una estimación de la nueva media del proceso después del corrimiento. Ésta puede calcularse con

$$\hat{\mu} = \begin{cases} \mu_0 + K + \dfrac{s_H(i)}{n_H}, \text{ si } s_H(i) > H \\ \mu_0 - K - \dfrac{s_L(i)}{n_L}, \text{ si } s_L(i) > H \end{cases} \tag{15-21}$$

También es conveniente presentar una representación gráfica de las CUSUM tabulares, a las que en ocasiones se les llama cartas de estado CUSUM. Se construyen graficando $s_H(i)$ y $s_L(i)$ contra el número de muestra. En la figura 15-20 se muestra la carta de estado CUSUM para los datos del ejemplo 15-6. Cada barra vertical representa el valor de $s_H(i)$ y $s_L(i)$ en el periodo i. Con el intervalo de decisión graficado en la carta, la carta de estado CUSUM es parecida a una carta de control de Shewhart. En la carta de estado CUSUM se han graficado también los estadísticos muestrales x_i para cada periodo utilizando puntos rellenos. Muchas veces esto ayuda al usuario de la carta de control a visualizar el desempeño real del proceso que ha llevado a un valor particular de la CUSUM.

La CUSUM tabular se diseña eligiendo valores para el valor de referencia K y para el intervalo de decisión H. Se recomienda que estos parámetros se seleccionen de tal modo que produzcan valores adecuados de la longitud de la corrida promedio. Se han realizado muchos estudios analíticos del desempeño de la ARL de una CUSUM. Con base en estos estudios, es posible ofrecer algunas recomendaciones generales para seleccionar H y K. Se definen $H = h\sigma_{\bar{X}}$ y $K = k\sigma_{\bar{X}}$, donde $\sigma_{\bar{X}}$ es la desviación estándar de la variable muestral usada para formar la CUSUM (si $n = 1, \sigma_{\bar{X}} = \sigma_X$). Al utilizar $h = 4$ o $h = 5$ y $k = 1/2$, por lo general se llegará a una CUSUM que

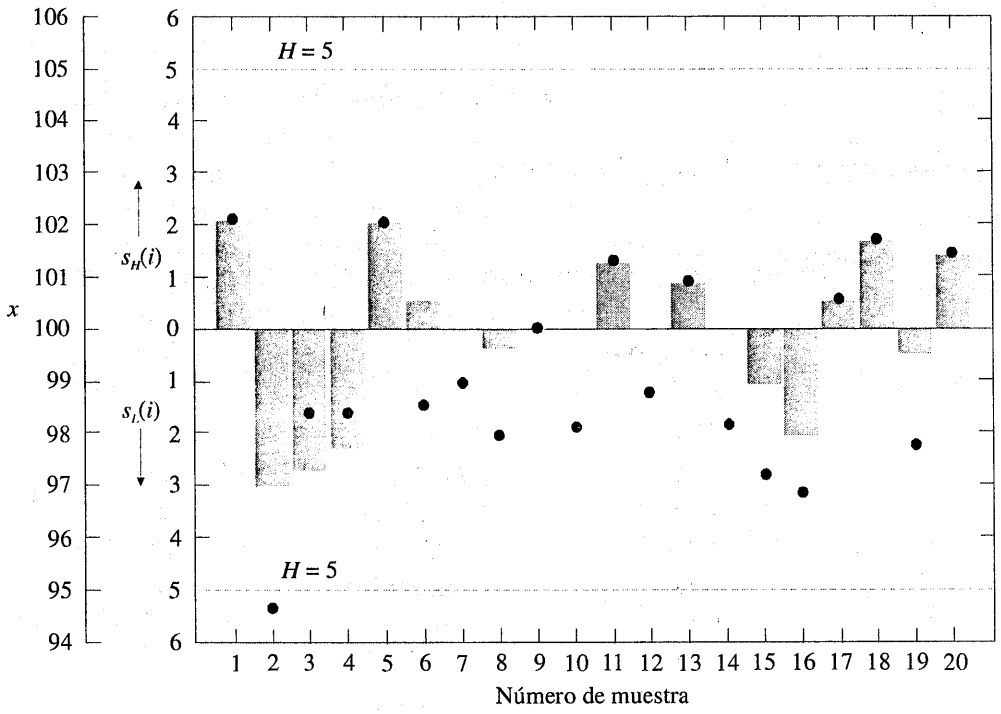


Figura 15-20 La carta del estado CUSUM para el ejemplo 15-6.

tenga propiedades adecuadas de la ARL contra un corrimiento aproximado de 1 $\sigma_{\bar{X}}$ (o $1\sigma_X$) en la media del proceso. Si el interés se encuentra en corrimientos mucho más grandes o pequeños, se hace $k = \delta/2$, donde δ es el tamaño del cambio en unidades de desviación estándar. Algunos autores prefieren usar una variable estandarizada $y_i = (\bar{x}_i - \mu_0)/\sigma_{\bar{X}}$ como base de la CUSUM. En este caso, las ecuaciones 15-19 y 15-20 quedan como

$$s_H(i) = \text{máx}[0, y_i - K + s_H(i-1)]$$

y

$$s_L(i) = \text{máx}[0, K - y_1 + s_L(i-1)]$$

Para este esquema, se seleccionaría $K = 1/2$ y $H = 4$ o $H = 5$.

Para ilustrar qué tan bien funcionan las recomendaciones de $h = 4$ o $h = 5$ con $k = 1/2$, considérense las longitudes de la corrida promedio de la tabla 15-9. Obsérvese que un corrimiento de $1\sigma_X$ se detectaría en 8.38 muestras (con $k = 1/2$ y $h = 4$) o bien en 10.4 muestras (con $k = 1/2$ y $h = 5$). En comparación, la tabla 15-1 indica que en una carta $\bar{X}$ se requerirían aproximadamente 43.9 muestras, en promedio, para detectar este corrimiento.

Estas reglas de diseño se usaron para la CUSUM del ejemplo 15-6. Se supuso que la desviación estándar del proceso es $\sigma = 2$. (Este es un valor razonable; ver el ejemplo 15-2.) Entonces con $k = 1/2$ y $h = 5$, se usaría

$$K = k\sigma = 1/2(2) = 1$$

y

$$H = h\sigma = 5(2) = 10$$

en el procedimiento de la CUSUM tabular.

Tabla 15-9 Longitudes de la serie promedio para una carta de control de CUSUM con $K = 1/2$

Cambio en la media (múltiplos de $\sigma_{\bar{X}}$)	$h = 4$	$h = 5$
0	168	465
0.25	74.2	139
0.50	26.6	38.0
0.75	13.3	17.0
1.00	8.38	10.4
1.50	4.75	5.75
2.00	3.34	4.01
2.50	2.62	3.11
3.00	2.19	2.57
4.00	1.71	2.01

Por último, cabe hacer notar que procedimientos complementarios tales como las reglas Western Electric no pueden aplicarse con seguridad a la CUSUM, debido a que los valores consecutivos de $S_H(i)$ y $S_L(i)$ no son independientes. De hecho, la CUSUM puede considerarse como un promedio ponderado, donde las ponderaciones son estocásticas o aleatorias. En realidad, los valores CUSUM tienen una correlación alta, ocasionando en consecuencia que las reglas Western Electric produzcan demasiadas falsas alarmas.

EJERCICIOS DE LA SECCIÓN 15-10

15-36. La pureza de un producto químico se mide cada dos horas. Los resultados de 20 mediciones consecutivas son los siguientes:

Muestra	Pureza	Muestra	Pureza
1	89.11	11	88.55
2	90.59	12	90.43
3	91.03	13	91.04
4	89.46	14	88.17
5	89.78	15	91.23
6	90.05	16	90.92
7	90.63	17	88.86
8	90.75	18	90.87
9	89.65	19	90.73
10	90.15	20	89.78

a) Construya una carta de control CUSUM para este proceso. Use $\sigma = 0.8$ para establecer el procedimiento, y suponga que el objetivo deseado del proceso es 90. ¿El proceso parece estar bajo control?

b) Suponga que las cinco observaciones siguientes son 90.75, 90.00, 91.15, 90.95 y 90.86. Aplique la CUSUM del inciso a) a estas nuevas observaciones. ¿Hay evidencia de que el proceso haya tenido un corrimiento fuera de control?

15-37. El diámetro de unos agujeros se mide en orden consecutivo con un sensor automático. Los resultados de la medición de 25 agujeros son los siguientes.

Muestra	Diámetro	Muestra	Diámetro
1	9.94	14	9.99
2	9.93	15	10.12
3	10.09	16	9.81
4	9.98	17	9.73
5	10.11	18	10.14
6	9.99	19	9.96
7	10.11	20	10.06
8	9.84	21	10.11
9	9.82	22	9.95
10	10.38	23	9.92
11	9.99	24	10.09
12	10.41	25	9.85
13	10.36		

a) Estime la desviación estándar del proceso.

b) Establezca un procedimiento de control CUSUM, suponiendo que el diámetro objetivo es 10.0 mm. ¿El proceso parece estar operando en un estado de control estadístico en el nivel objetivo deseado?

15-38. La concentración de un producto químico se mide tomando cuatro muestras de cada lote de material. En la tabla siguiente se muestra la concentración promedio de estas mediciones para los últimos 20 lotes.

Muestra	Concentración	Muestra	Concentración
1	104.5	11	95.4
2	99.9	12	94.5
3	106.7	13	104.5
4	105.2	14	99.7
5	94.8	15	97.7
6	94.6	16	97.0
7	104.4	17	95.8
8	99.4	18	97.4
9	100.3	19	99.0
10	100.3	20	102.6

a) Suponga que la desviación estándar del proceso es $\sigma = 8$ y que el valor objetivo de la concentración para este proceso es 100. Diseñe un esquema CUSUM para el proceso. ¿El proceso parece estar bajo control en el objetivo?

b) ¿Cuántos lotes se esperaría producir con una concentración fuera del objetivo antes de que la carta de control CUSUM lo detectara si la concentración tiene un corrimiento a 104? Use la tabla 15-9.

15-39. Considérese una CUSUM estandarizada con $H = 5$ y $K = 1/2$. Se toman muestras del proceso cada dos horas. El valor objetivo del proceso es $\mu_0 = 50$ y $\sigma = 2$. Use la tabla 15-9.

a) Si el tamaño de la muestra es $n = 1$, ¿cuántas muestras se necesitarían para detectar un corrimiento en la media del proceso a $\mu = 51$ en promedio?

b) Si el tamaño de la muestra se aumenta a $n = 4$, ¿en qué forma afecta esto la longitud de la corrida promedio para detectar el corrimiento a $\mu = 51$ que se determinó en el inciso a)?

15-40. Un proceso tiene un objetivo de $\mu_0 = 100$ y una desviación estándar de $\sigma = 4$. Se toman muestras de tamaño $n = 1$ cada dos horas. Use la tabla 15-9.

a) Suponga que la media del proceso se corre a $\mu = 102$. ¿Cuántas horas de producción transcurrirán antes de que el corrimiento del proceso sea detectado por una CUSUM estandarizada con $H = 5$ y $K = 1/2$?

b) Es importante detectar con mayor rapidez el corrimiento definido en el inciso a). Se hace una propuesta para reducir la frecuencia de muestreo a 0.5 horas. ¿En qué forma afecta esto el procedimiento de control CUSUM? ¿Con cuánta más rapidez se detectará el corrimiento?

c) Suponga que se adopta el intervalo de muestreo de 0.5 horas del inciso b). ¿Con qué frecuencia ocurrirán falsas alarmas con este nuevo intervalo de muestreo? ¿Con qué frecuencia ocurrían con el intervalo anterior de dos horas?

d) Se hace una propuesta para aumentar el tamaño de la muestra a $n = 4$ y conservar el intervalo de muestreo de dos horas. ¿Cómo se compara esta propuesta en términos del tiempo de detección promedio con la sugerencia de disminuir el intervalo de muestreo a 0.5 horas?

15-11 OTRAS HERRAMIENTAS DEL CONTROL ESTADÍSTICO DE PROCESOS PARA RESOLVER PROBLEMAS

Aun cuando la carta de control es una herramienta muy poderosa para investigar las causas de la variación de un proceso, es más efectiva cuando se usa con otras herramientas del CEP para resolver problemas. En esta sección se ilustrarán algunas de estas herramientas, utilizando los datos de los defectos en las tarjetas de circuitos impresos del ejemplo 15-4.

En la figura 15-16 se muestra una carta U para el número de defectos en muestras de cinco tarjetas de circuitos impresos. La carta indica control estadístico, pero es necesario reducir el número de defectos. El número promedio de defectos por tarjeta es $8/5 = 1.6$, y este nivel de defectos requeriría un reprocesamiento considerable.

El primer paso para resolver este problema es construir un **diagrama de Pareto** de los tipos de defectos individuales. El diagrama de Pareto, ilustrado en la figura 15-21, indica que la soldadura insuficiente y las bolas de soldadura son los defectos que ocurren con mayor frecuencia, representando $(109/160)100 = 68\%$ de los defectos observados. Por otra parte, las cinco primeras categorías de defectos del diagrama de Pareto son en su totalidad de defectos relacio-

nadas con la soldadura. Esto apunta hacia el proceso de soldadura líquida como una oportunidad potencial de mejoramiento.

Para mejorar el proceso de soldadura líquida, se reúne un equipo integrado por el operador de la máquina de soldadura líquida, el supervisor del taller, el ingeniero de manufactura responsable del proceso y un ingeniero de calidad para estudiar las causas potenciales de los defectos de soldadura. Realizan una sesión de lluvia de ideas y elaboran el **diagrama de causa y efecto** de la figura 15-22. El diagrama de causa y efecto es de uso común para representar las diversas causas de los defectos en los productos y sus interrelaciones. Son útiles para resumir el conocimiento que se tiene del proceso.

Como resultado de la sesión de lluvia de ideas, el equipo hace la identificación tentativa de las siguientes variables como factores que influirían potencialmente en la generación de defectos de soldadura:

1. Peso específico del flujo
2. Temperatura de la soldadura
3. Velocidad de la transportadora
4. Ángulo de la transportadora
5. Altura de la onda de soldadura
6. Temperatura de precalentamiento
7. Método para cargar la paleta

Podría usarse un **experimento diseñado** estadísticamente para investigar el efecto de estas siete variables sobre los defectos de soldadura.

Además, el equipo construyó un diagrama de la **concentración de defectos** para el producto. Un diagrama de la concentración de defectos es tan sólo un trazo o dibujo del producto, con

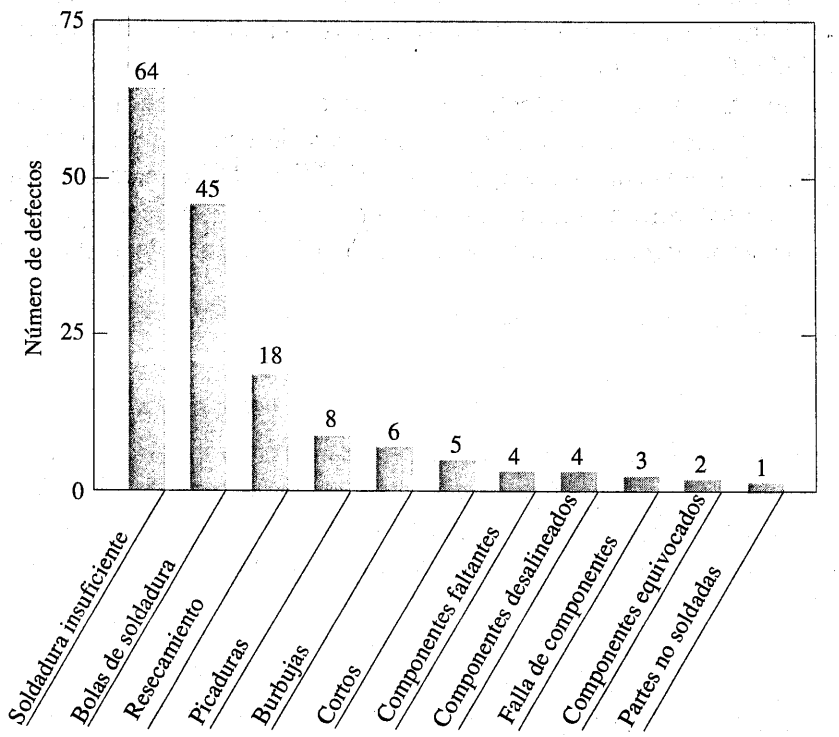


Figura 15-21 Diagrama de Pareto para los defectos de las tarjetas de circuitos impresos.

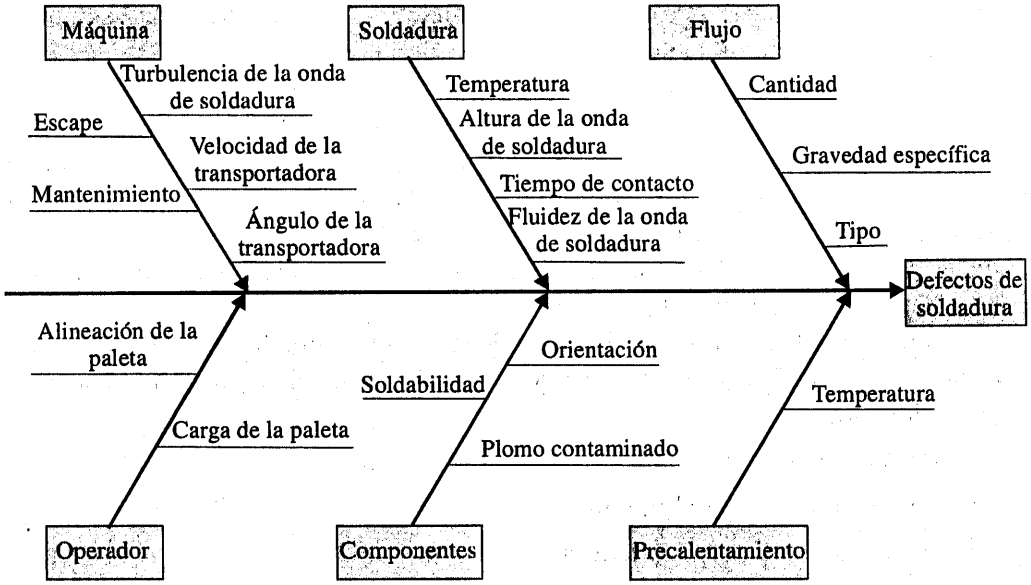


Figura 15-22 Diagrama de causa y efecto para el proceso de soldadura líquida de tarjetas de circuitos impresos.

los defectos que ocurren con mayor frecuencia indicados en la pieza. Este diagrama se usa para determinar si los defectos ocurren en el mismo lugar de la pieza. En la figura 15-23 se muestra el diagrama de la concentración de defectos para las tarjetas de circuitos impresos. Este diagrama indica que la mayoría de los defectos de soldadura insuficiente están cerca del extremo frontal de la tarjeta, donde hace contacto inicial con la onda de soldadura. Una investigación adicional indicó que una de las paletas usadas para transportar las tarjetas a través de la onda de soldadura estaba doblada, ocasionando que el extremo frontal de la tarjeta tuviera un contacto deficiente con la onda de soldadura.

Después de reemplazar la paleta defectuosa, se utilizó un experimento diseñado para investigar las siete variables analizadas arriba. Los resultados de este experimento indicaron que varios de estos factores influían y podían ajustarse para reducir los defectos de soldadura. Después de que se implementaron los resultados del experimento, el porcentaje de las juntas de soldadura que requirieron reprocesamiento se redujo de 1% a menos de 100 partes por millón (0.01%).

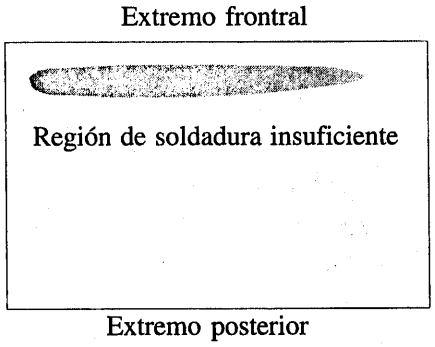


Figura 15-23 Diagrama de la concentración de los defectos para una tarjeta de circuitos impresos.

15-12 IMPLEMENTACIÓN DEL CONTROL ESTADÍSTICO DE PROCESOS

Los métodos del control estadístico de procesos (CEP) pueden redituar enormes beneficios a las compañías que puedan implementarlos con éxito. Aun cuando el CEP da la impresión de ser una colección de herramientas para la resolución de problemas basadas en la estadística, su uso exitoso es algo más que el simple aprendizaje y aplicación de estas herramientas. La participación y el compromiso de la administración con el proceso de mejoramiento de calidad es el componente de mayor peso en el éxito potencial del CEP. La administración constituye un modelo a seguir, y otras áreas de la organización se volverán hacia la administración en busca de guía y ejemplo. El trabajo en equipo también es importante, ya que por lo general es difícil que una sola persona introduzca mejoras en el proceso. La mayoría de las "siete magníficas" herramientas para la resolución de problemas pueden contribuir para constituir un equipo de mejoras, incluyendo los diagramas de causa y efecto, los diagramas de Pareto y los diagramas de la concentración de los defectos. Las herramientas básicas para resolver problemas del CEP deberán conocerse y utilizarse ampliamente en la organización completa. La capacitación continua en CEP y en el mejoramiento de calidad es necesaria para lograr este conocimiento generalizado de las herramientas.

El objetivo de un programa de mejoramiento de la calidad basado en el CEP es el mejoramiento continuo sobre una base semanal, trimestral y anual. El CEP no es un programa que deba aplicarse cuando el negocio pasa por problemas y que se abandone después. El mejoramiento de la calidad debe convertirse en parte de la cultura de la organización.

La carta de control es una herramienta importante en el mejoramiento de la calidad. Los procesos no operan en un estado bajo control de manera natural, y el uso de cartas de control es un paso importante que debe tomarse al iniciar un programa de CEP para eliminar las causas asignables, reducir la variabilidad del proceso y estabilizar el desempeño del proceso. A fin de mejorar la calidad y la productividad, debe empezarse por llevar la administración con base en hechos y datos, y no confiar únicamente en el criterio personal. Las cartas de control son una parte importante de este cambio en el enfoque administrativo.

Al implementar un programa de CEP a nivel de la compañía, los autores han encontrado que los siguientes elementos suelen estar presentes en todos los esfuerzos exitosos:

1. Liderazgo administrativo
2. Enfoque de equipos
3. Capacitación de los empleados de todos los niveles
4. Énfasis en el mejoramiento continuo
5. Mecanismos para identificar el éxito

Hay que insistir en la importancia que tiene el liderazgo administrativo y el enfoque de equipos. El mejoramiento de la calidad exitoso es una actividad impulsada por la administración "de arriba hacia abajo". También es importante medir los avances y el éxito, y difundir el conocimiento de este éxito en toda la organización. Cuando se comunican los mejoramientos exitosos en toda la compañía, esto puede proporcionar motivación e incentivo para mejorar otros procesos y para hacer del mejoramiento continuo una parte normal de la manera de hacer negocios.

La filosofía de W. Edwards Deming ofrece un importante marco de referencia para implementar el mejoramiento de la calidad y la productividad. La filosofía de Deming se resume en sus 14 puntos para la administración. La adhesión a estos principios de administración ha

sido factor importante en el éxito industrial de Japón y sigue siendo el catalizador de los esfuerzos de mejoramiento de la calidad y la productividad en aquella nación. Esta filosofía también se está difundiendo rápidamente en Occidente. Los 14 puntos de Deming son los siguientes.

1. **Ser constante en las metas enfocada en el mejoramiento de los productos y los servicios.** Intentar de manera constante el mejoramiento del diseño de los productos y de su rendimiento. La inversión en investigación, desarrollo e innovación rendirá frutos a largo plazo para la organización.
2. **Adoptar una nueva filosofía de rechazar la mano de obra deficiente, los productos defectuosos o los malos servicios.** Cuesta lo mismo producir una unidad defectuosa que producir una unidad satisfactoria (y en ocasiones más). El costo de tratar con productos de desecho, tareas de reprocesamiento y otras pérdidas derivadas de las unidades defectuosas constituye un enorme consumo de los recursos de una compañía.
3. **No confiar en la inspección en masa para "controlar" la calidad.** Lo único que puede hacer la inspección es descartar las unidades defectuosas, y en este punto es demasiado tarde, porque ya se ha pagado por producirlas. La inspección ocurre demasiado tarde en el proceso, es costosa y con frecuencia ineficaz. La calidad es el resultado de la prevención de unidades defectuosas a través del mejoramiento del proceso, no de la inspección.
4. **No hacer contratos con proveedores únicamente con base en el precio, sino considerar también la calidad.** El precio sólo es una medición importante del producto del proveedor si se considera en relación con una medida de la calidad. En otras palabras, debe tomarse en cuenta el costo total del artículo, no sólo el precio de compra. Cuando se toma en consideración la calidad, muchas veces el postor más bajo no es el proveedor con el costo más bajo. Deberá darse preferencia a los proveedores que usan métodos modernos de mejoramiento de calidad en sus negocios y que pueden demostrar el control y la capacidad del proceso.
5. **Enfocarse en el mejoramiento continuo.** Intentar de manera constante mejorar la producción y el sistema de servicios. Comprometer a la fuerza de trabajo en estas actividades y hacer uso de los métodos estadísticos, en particular las herramientas del CEP para la solución de problemas analizadas en la sección anterior.
6. **Poner en práctica métodos de capacitación modernos e invertir en la capacitación de todos los empleados.** Todos deberán recibir capacitación en los aspectos técnicos de su trabajo, así como en los métodos modernos de mejoramiento de la calidad y la productividad. La capacitación deberá estimular a todos los empleados a poner cotidianamente en práctica estos métodos.
7. **Poner en práctica métodos de supervisión modernos.** La supervisión no deberá consistir simplemente en la vigilancia pasiva de los trabajadores, sino que deberá enfocarse en ayudar a los empleados a mejorar el sistema en el cual trabajan. La primera meta de la supervisión deberá ser mejorar el sistema de trabajo y el producto.
8. **Erradicar el temor.** Muchos trabajadores temen hacer preguntas, reportar problemas o señalar condiciones que son barreras para la calidad y la producción eficaz. En muchas organizaciones las pérdidas económicas asociadas con el temor son cuantiosas; la administración es la única que puede eliminar el temor.

9. **Derribar las barreras entre las áreas funcionales del negocio**. El trabajo en equipo entre las diferentes unidades organizacionales es esencial para que tenga lugar el mejoramiento efectivo de la calidad y la productividad.
10. **Eliminar los objetivos, los lemas y las metas numéricas para la fuerza laboral**. Un objetivo tal como "cero defectos" es inútil sin un plan en cuanto a la manera de alcanzar este objetivo. De hecho, estos lemas y "programas" por lo general resultan contraproducentes. Trabajar para mejorar el sistema y proporcionar información al respecto.
11. **Eliminar las cuotas numéricas y los estándares de trabajo**. Estos estándares se han establecido históricamente sin considerar la calidad. Los estándares de trabajo con frecuencia son síntomas de la incapacidad de la administración para entender el proceso laboral y para proporcionar un sistema administrativo eficaz enfocado en el mejoramiento de este proceso.
12. **Eliminar las barreras que desestimulan a los empleados a hacer su trabajo**. La administración debe escuchar las sugerencias, comentarios y quejas de los empleados. La persona que está realizando el trabajo es la que lo conoce mejor, y por lo general tiene ideas valiosas acerca de cómo conseguir que el proceso opere con mayor eficiencia. La fuerza de trabajo es un participante importante del negocio y no sólo el oponente en la negociación de contratos colectivos.
13. **Instituir un programa continuo de capacitación y educación para todos los empleados**. La capacitación en técnicas estadísticas, sencillas, pero poderosas, debe ser obligatoria para todos los empleados. El uso de las herramientas de CEP básicas para la solución de problemas, en particular de la carta de control, deberá difundirse en todo el negocio. Cuando estas cartas se generalicen, y cuando los empleados entiendan sus aplicaciones, será más probable que éstos busquen las causas de la calidad deficiente y que identifiquen las mejoras del proceso. La capacitación es una forma de hacer que todos sean socios en el proceso de mejoramiento de calidad.
14. **Crear una estructura en la gerencia general que se aboque enérgicamente a los 13 primeros puntos**.

Cuando se leen los 14 puntos de Deming, se observan dos cosas. Primera, hay un fuerte énfasis en el cambio. Segunda, el papel de la administración para guiar este proceso de cambio es de importancia dominante. Pero, ¿qué deberá cambiarse y cómo deberá iniciarse este proceso de cambio? Por ejemplo, si quiere mejorarse el rendimiento de un proceso de manufactura de semiconductores, ¿qué debe hacerse? Es en esta área donde los métodos estadísticos entran en juego con mayor frecuencia. Para mejorar el proceso de los semiconductores, debe determinarse cuáles son los factores controlables del proceso que influyen en el número de unidades defectuosas producidas. Para responder a esta pregunta, deben recabarse datos sobre el proceso y ver cómo reacciona el sistema a los cambios en las variables del proceso. Los métodos estadísticos, incluyendo el CEP y las técnicas de diseño experimental de este libro, pueden contribuir a este conocimiento.

Ejercicios complementarios

15-41. El diámetro de los fusibles de clavija usados en una aplicación de una turbina de avión es una característica de la calidad importante. En la página siguiente se describen veinticinco muestras de tres pernos cada una (en mm):

a) Establezca las cartas $\bar{X}$ y R para este proceso. De ser necesario, revise los límites

Número de Muestra	Diámetro		
1	64.030	64.002	64.019
2	63.995	63.992	64.001
3	63.988	64.024	64.021
4	64.002	63.996	63.993
5	63.992	64.007	64.015
6	64.009	63.994	63.997
7	63.995	64.006	63.994
8	63.985	64.003	63.993
9	64.008	63.995	64.009
10	63.998	74.000	63.990
11	63.994	63.998	63.994
12	64.004	64.000	64.007
13	63.983	64.002	63.998
14	64.006	63.967	63.994
15	64.012	64.014	63.998
16	64.000	63.984	64.005
17	63.994	64.012	63.986
18	64.006	64.010	64.018
19	63.984	64.002	64.003
20	64.000	64.010	64.013
21	63.988	64.001	64.009
22	64.004	63.999	63.990
23	64.010	63.989	63.990
24	64.015	64.008	63.993
25	63.982	63.984	63.995

para que ninguna de las observaciones esté fuera de control.

b) Estime la media y la desviación estándar del proceso.

c) Suponga que las especificaciones del proceso están en 64 ± 0.02. Calcule una estimación del *ICP*. ¿El proceso satisface un nivel de capacidad mínimo de $ICP \geq 1.33$?

d) Calcule una estimación de ICP_k. Use este índice para sacar conclusiones acerca de la capacidad del proceso.

e) Para hacer que este proceso sea seis-sigma, la varianza σ^2 tendría que reducirse de tal modo que $ICP_k = 2.0$. ¿Cuál deberá ser este nuevo valor de la varianza?

f) Suponga que la media se corre a 64.005. ¿Cuál es la probabilidad de que este corrimiento se detectará en la siguiente muestra? ¿Cuál es la ARL después del corrimiento?

15-42. Las botellas de plástico para un detergente líquido para lavar ropa se hacen por moldeo de soplado. Se inspeccionan 20 muestras de $n = 100$ botellas en el orden en el tiempo en que se producen, y se reporta la proporción defectuosa de cada muestra. Los datos se muestran a continuación:

Muestra	Proporción Defectuosa
1	0.12
2	0.15
3	0.18
4	0.10
5	0.12
6	0.11
7	0.05
8	0.09
9	0.13
10	0.13
11	0.10
12	0.07
13	0.12
14	0.08
15	0.09
16	0.15
17	0.10
18	0.06
19	0.12
20	0.13

a) Construya una carta *P* para este proceso. ¿El proceso está bajo control estadístico?

b) Suponga que en vez de $n = 100$, se tiene $n = 200$. Use los datos dados para construir una carta *P* para este proceso. Revise los límites de control en caso necesario.

c) Compare los límites de control de las cartas *P* de los incisos a) y b). Explique por qué difieren. Además, explique por qué la determinación del control estadístico difiere para los dos tamaños de *n*.

15-43. Unas cubiertas para computadoras personales se fabrican por moldeo por inyección. Se toman periódicamente muestras de cinco cubiertas del proceso, y se registra el núme-

ro de defectos. En seguida se presentan los datos de 25 muestras.

Muestra	Núm. de defectos
1	3
2	2
3	0
4	1
5	4
6	3
7	2
8	4
9	1
10	0
11	2
12	3
13	2
14	8
15	0
16	2
17	4
18	3
19	5
20	0
21	2
22	1
23	9
24	3
25	2

a) Utilizando todos los datos, encuentre los límites de control de prueba para esta carta U del proceso.

b) Use los límites de control de prueba del inciso a) para identificar los puntos fuera de control. De ser necesario, revise los límites de control.

c) Suponga que en lugar de muestras de 5 cubiertas, el tamaño de la muestra fue de 10. Repita los incisos a) y b). Explique en qué forma modifica este cambio las respuestas de los incisos a) y b).

15-44. Considérense los datos del ejercicio 15-43.

a) Utilizando todos los datos, encuentre los límites de control de prueba para una carta C para el proceso.

b) Use los límites de control de prueba del inciso a) para identificar los puntos fuera de control. De ser necesario, revise los límites de control.

c) Suponga que en lugar de muestras de 5 cubiertas, la muestra fue de 10 cubiertas. Repita los incisos a) y b). Explique en qué forma modifica este cambio las respuestas de los incisos a) y b).

15-45. Suponga que un proceso está bajo control y que se usa una carta $\bar{X}$ con un tamaño de la muestra de 4 para monitorear el proceso. De pronto ocurre un corrimiento en la media de 1.5σ.

a) Si se están usando límites de control 3- sigma en la carta $\bar{X}$ ¿cuál es la probabilidad de que este corrimiento permanezca sin ser detectado en tres muestras consecutivas?

b) Si se están usando límites de control 2-sigma en la carta $\bar{X}$, ¿cuál es la probabilidad de que este corrimiento permanezca sin ser detectado en tres muestras consecutivas?

c) Compare las respuestas de los incisos a) y b) y explique por qué difieren. Además, ¿qué límites se recomendaría usar y por qué?

15-46. Considere la carta de control para mediciones individuales con límites 3-sigma.

a) Suponga que ocurre un corrimiento de magnitud σ en la media del proceso. Verifique que la ARL para detectar el corrimiento es ARL = 43.9.

b) Encuentre la ARL para detectar un corrimiento de magnitud 2σ en la media del proceso.

c) Encuentre la ARL para detectar un corrimiento de magnitud 3σ en la media del proceso.

d) Compare las respuestas de los incisos a), b) y c) y explique por qué la ARL de detección disminuye cuando la magnitud del corrimiento se incrementa.

15-47. Considérese una carta de control para mediciones individuales, aplicado a un proceso químico continuo de 24 horas con observaciones tomadas cada hora.

a) Si la carta tiene límites 3-sigma, verifique que la ARL del estado bajo control es ARL = 370. ¿Cuántas falsas alarmas ocurrirían en cada mes de 30 días, en promedio, con esta carta?

b) Suponga que la carta tiene límites 2-sigma. ¿Esto reduce la ARL para detectar un corrimiento de magnitud σ en la media? (Recuérdese que la ARL para detectar este corrimiento con límites 3-sigma es 43.9.)

c) Encuentre la ARL bajo control si en la carta se usan límites 2-sigma. ¿Cuántas falsas alarmas ocurrirían cada mes con esta carta? ¿Es satisfactorio el desempeño de esta ARL bajo control? Explique su respuesta.

15-48. La profundidad de una ranura es una importante característica de la calidad de una pieza. Se toman muestras de tamaño $n = 5$ cada cuatro horas del proceso, y se resumen a continuación 20 muestras:

Muestra	$\bar{x}$	r
1	139.7	1.1
2	139.8	1.4
3	140.0	1.3
4	140.1	1.6
5	139.8	0.9
6	139.9	1.0
7	139.7	1.4
8	140.2	1.2
9	139.3	1.1
10	140.7	1.0
11	138.4	0.8
12	138.5	0.9
13	137.9	1.2
14	138.5	1.1
15	140.8	1.0
16	140.5	1.3
17	139.4	1.4
18	139.9	1.0
19	137.5	1.5
20	139.2	1.3

a) Utilizando todos los datos, encuentre los límites de control de prueba para las cartas $\bar{X}$ y R. ¿El proceso está bajo control?

b) Use los límites de control de prueba del inciso a) para identificar los puntos fuera de control. De ser necesario, revise los límites de control. Ahora, estime la desviación estándar del proceso.

c) Suponga que las especificaciones están en 140 ± 2. Utilizando los resultados del inciso b), ¿qué afirmaciones pueden hacerse acerca de la capacidad del proceso? Calcule estimaciones de los índices apropiados de la capacidad del proceso.

d) Para hacer este proceso "seis-sigma", la varianza σ^2 tendría que reducirse de tal modo que $ICP_k = 2.0$. ¿Cuál deberá ser este nuevo valor de la varianza?

e) Suponga que la media se corre a 139.7. ¿Cuál es la probabilidad de que este corrimiento se detecte en la siguiente muestra? ¿Cuál es la ARL después del corrimiento?

15-49. Un proceso se controla con una carta P utilizando muestras de tamaño 100. La línea central de la carta es 0.05.

a) ¿Cuál es la probabilidad de que la carta de control detecte un corrimiento a 0.08 en la primera muestras después del corrimiento?

b) ¿Cuál es la probabilidad de que la carta de control no detecte un corrimiento a 0.08 en la primera muestra después del corrimiento, pero lo detecte en la segunda muestra?

c) Suponga que en lugar de un corrimiento a 0.08 en la media, ésta se corre a 0.10. Repita los incisos a) y b).

d) Compare las respuestas para un corrimiento a 0.08 y para un corrimiento a 0.10. Explique por qué difieren. Además, explique por qué es más sencillo detectar un corrimiento a 0.10.

15-50. Suponga que se sabe que el número promedio de defectos en una unidad es 8. Si el número de defectos promedio en una unidad se corre a 16, ¿cuál es la probabilidad de

que será detectado por una carta U en la primera muestra después del corrimiento,
a) si el tamaño de la muestra es $n = 4$?
b) si el tamaño de la muestra es $n = 10$?
Use una aproximación normal para U.

15-51. Suponga que se sabe que el número de defectos promedio en una unidad es 10. Si el número de defectos promedio en una unidad se corre a 14, ¿cuál es la probabilidad de que será detectado por una carta U en la primera muestra después del corrimiento,
a) si el tamaño de la muestra es $n = 1$?
b) si el tamaño de la muestra es $n = 5$?
Use una aproximación normal para U.

15-52. Suponga que se usa una carta de control $\bar{X}$ con límites 2-sigma para controlar un proceso. Encuentre la probabilidad de que se produzca una falsa señal de fuera de control en la siguiente muestra. Compare esto con la probabilidad correspondiente para la carta con límites 3-sigma y comente los resultados. Comente asimismo cuándo sería preferible usar límites 2-sigma en lugar de límites 3-sigma.

15-53. Considérese una carta de control $\bar{X}$ con límites de control k-sigma. Desarrolle una expresión general para la probabilidad de que un punto se localizará fuera de los límites de control cuando la media del proceso se ha corrido δ unidades de la línea central.

15-54. Considérese la carta de control $\bar{X}$ con límites 2-sigma del ejercicio 15-47.
a) Encuentre la probabilidad de que no haya ninguna señal en la primera muestra pero que haya una señal en la segunda.
b) ¿Cuál es la probabilidad de que no habrá una señal en tres muestras?

15-55. Suponga que un proceso tiene un $ICP = 2$, pero la media está exactamente tres desviaciones estándar arriba del límite superior de las especificaciones. ¿Cuál es la probabilidad de fabricar un producto fuera de los límites de las especificaciones?

15-56. Considérense los datos de la medición de la dureza del ejercicio 15-8. Establezca un esquema CUSUM para este proceso usando $\mu = 50$ y $\sigma = 2$, de tal modo que $K = 1$ y $H = 10$. ¿El proceso está bajo control?

15-57. Considere los datos del ejercicio 15-9. Establezca un esquema CUSUM para este proceso suponiendo que $\mu = 80$ es el objetivo del proceso. Explique cómo se determinaron la estimación de σ y los parámetros CUSUM K y H.

15-58. Considere de nuevo los datos del ejercicio 15-11. Construya una carta de control CUSUM para este proceso, usando $\mu_0 = 500$ como el objetivo del proceso. Explique cómo se determinaron la estimación de σ y los parámetros CUSUM H y K.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

15-59. Suponga que un proceso está bajo control, y que en la carta $\bar{X}$ se están usando límites de control 3-sigma. Sea que la media se corre en 1.5σ. ¿Cuál es la probabilidad de que este corrimiento permanezca sin ser detectado en tres muestras consecutivas? ¿Cuál sería esta probabilidad si se usaran límites de control 2-sigma? El tamaño de la muestra es 4.

15-60. Suponga que se usa una carta $\bar{X}$ para controlar un proceso con una distribución normal y que se toman muestras de tamaño n cada h horas y se grafican en la carta, la cual tiene límites k-sigma.
a) Encuentre una expresión general para el número esperado de muestras y de tiempo que se tomarán hasta que se genere una señal de acción.

EJERCICIOS PARA DESARROLLAR EL INTELECTO

b) Suponga que la media del proceso se corre a un estado fuera de control, por ejemplo $\mu_1 = \mu_0 + \delta\sigma$. Encuentre una expresión para el número esperado de muestras que se tomarán hasta que se genere una falsa acción.

c) Evalúe la ARL bajo control, para $k = 3$. ¿Cómo cambia esto si $k = 2$? ¿Qué opina usted del uso de límites 2-sigma en la práctica?

d) Evalúe la ARL fuera de control para un cambio de 1 sigma, dado que $n = 5$.

15-61. Suponga que se usa una carta P con línea central en $\bar{p}$ y límites de control k-sigma para controlar un proceso. Hay una proporción defectuosa crítica p_c que debe detectarse con una probabilidad de 0.50 en la primera muestra después del corrimiento a este estado. Deduzca una fórmula general para el tamaño de la muestra que deberá usarse en esta carta.

15-62. Suponga que se usa una carta P con línea central en $\bar{p}$ y límites de control k-sigma para controlar un proceso. ¿Cuál es el tamaño de la muestra menor que puede usarse en esta carta de control para asegurar que el límite de control inferior es positivo?

15-63. Un proceso se controla con una carta P utilizando muestras de tamaño 100. La línea central de la carta es 0.05. ¿Cuál es la probabilidad de que la carta de control detecte un corrimiento a 0.08 en la primera muestra después del corrimiento? ¿Cuál es la probabilidad de que el corrimiento se detecte al menos en la tercera muestra después del corrimiento?

15-64. Considérese un proceso en el que las especificaciones de una característica de la calidad son 100 ± 15. Se sabe que es 5 la desviación estándar de esta característica de la calidad que tiene una distribución normal. ¿Dónde se centraría el proceso para minimizar la proporción defectuosa producida? Suponga ahora que la media se corre a 105 y que se está usando un tamaño de la muestra de 4 en una carta $\bar{X}$. ¿Cuál es la probabilidad de que un cambio como éste se detectará en la primera muestra después del corrimiento? ¿Cuál es el número promedio de muestras hasta que ocurre un punto fuera de control? Compare este resultado con el número promedio de observaciones hasta que ocurre una unidad defectuosa (suponga la normalidad).

15-65. La carta de control *NP*. Una alternativa de la carta de control para una fracción defectuosa es una carta de control basada en el número de unidades defectuosas, o carta de control *NP*. La carta tiene la línea central en $n\bar{p}$, y los límites de control son

$$LCS = n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$$

$$LCI = n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})}$$

y el número de unidades defectuosas de cada muestra se grafica en la carta.

a) Verifique que los límites de control dados arriba son correctos.

b) Aplique esta carta de control a los datos del ejemplo 15-3.

c) ¿Esta carta proporcionará siempre resultados que son equivalentes a la carta P ordinaria?

15-66. La carta de control EWMA. El promedio móvil ponderado exponencialmente (o EWMA, por sus siglas en inglés) se define como:

$$Z_t = \lambda\bar{X}_t + (1-\lambda)\bar{Z}_{t-1}$$

donde $0 < \lambda \leq 1$, y el valor inicial del EWMA en el tiempo $t = 0$ es $Z_0 = \mu_0$ (el objetivo del proceso). Una carta de control EWMA se construye graficando los valores Z_t en una carta con línea central en μ_0 y los límites de control apropiados.

a) Verifique que $E(Z_t) = \mu_0$

EJERCICIOS PARA DESARROLLAR EL INTELECTO

b) Sea $\sigma^2_{Z_t} = V(Z_t)$, y demuestre que

$$\sigma^2_{Z_t} = \frac{\sigma^2}{n}\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2t}\right]$$

c) Use el resultado del inciso b) para determinar los límites de control de la carta EWMA.

d) Cuando $\lambda \to 1$, la carta de control EWMA deberá comportarse como una carta $\bar{X}$ de Shewhart ordinaria. ¿Está usted de acuerdo con esta afirmación? ¿Por qué?

e) Cuando $\lambda \to 1$, la carta de control EWMA deberá comportarse como una CUSUM. Ofrezca un argumento de por qué esto es así.

f) Aplique este procedimiento a los datos del ejemplo 15-2.

15-67. La carta de control estandarizada. Considérese la carta P con los límites de control 3-sigma usuales. Suponga que se define una nueva variable:

$$Z_i = \frac{\hat{P}_i - \bar{P}}{\sqrt{\dfrac{\bar{P}(1-\bar{P})}{n}}}$$

como la cantidad que se graficará en una carta de control. Se propone que esta nueva carta tenga línea central en 0 con los límites de control superior e inferior en ±3. Verifique que esta carta de control estandarizada será equivalente a la carta p original.

15-68. Tamaños desiguales de las muestras. Una aplicación de la carta de control estandarizada introducida en el ejercicio 15-67 es permitir el uso de tamaños desiguales de las muestras en la carta de control. Proporcione detalles respecto de cómo se implementaría este procedimiento e ilustrarlos utilizando los siguientes datos:

Muestra, i	1	2	3	4	5	6	7	8	9	10
n_i	20	25	20	25	50	30	25	25	25	20
p_i	0.2	0.16	0.25	0.08	0.3	0.1	0.12	0.16	0.12	0.15

APÉNDICES

APÉNDICE A. TABLAS Y CARTAS ESTADÍSTICAS

Tabla I Resumen de distribuciones de probabilidad comunes A-2
Tabla II Distribución normal estándar acumulada A-3
Tabla III Puntos porcentuales $\chi^2_{\alpha,\nu}$, de la distribución ji-cuadrada A-5
Tabla IV Puntos porcentuales $t_{\alpha,\nu}$ de la distribución t A-6
Tabla V Puntos porcentuales f_{α,ν_1,ν_2} de la distribución F A-7
Carta VI Curvas de operación característica A-12
Carta VII Curvas de operación característica para el análisis de varianza del modelo con efectos fijos A-21
Carta VIII Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios A-25
Tabla IX Valores críticos para la prueba del signo A-29
Tabla X Valores críticos para la prueba del rango con signo de Wilcoxon A-30
Tabla XI Valores críticos para la prueba de la suma de rangos de Wilcoxon A-31
Tabla XII Factores para construir cartas de control de variables A-33

APÉNDICE B. MATERIAL TÉCNICO COMPLEMENTARIO B-1

I. Técnicas de conteo B-1
II. Función generadora de momentos B-7
III. Funciones de variables aleatorias B-15
IV. Desarrollo de las distribuciones t y F B-24
V. Enfoque bayesiano de la estimación B-28
VI. Pruebas del cociente de verosimilitud B-32
VII. Factores aleatorios en experimentos factoriales B-35

APÉNDICE C. BIBLIOGRAFÍA C-1

APÉNDICE D. RESPUESTAS DE EJERCICIOS SELECCIONADOS D-1

Apéndice A

Tablas y cartas estadísticas

Tabla I Resumen de distribuciones de probabilidad comunes

Nombre	Función de densidad de probabilidad	Media	Varianza	Sección del libro
Discreta				
Uniforme	$\frac{1}{b-a}, a \le b$	$\frac{(b+a)}{2}$	$\frac{(b-a+1)^2-1}{12}$	4-5
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$, $x=0,1,\ldots,n, 0 \le p \le 1$	np	$np\,(1-p)$	4-6
Geométrica	$(1-p)^{x-1} p$, $x=1,2,\ldots, 0 \le p \le 1$	$1/p$	$(1-p)/p^2$	4-7.1
Binomial negativa	$\binom{x-1}{r-1}(1 \le p)^{x-r} p^r$ $x=r,r+1,r+2,\ldots,$ $0 \le p \le 1$	r/p	$r(1-p)/p^2$	4-7.2
Hipergeométrica	$\frac{\binom{K}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$ $x=\text{máx}(0,n-N+K), 1, \ldots$ $\text{mín}(K,n), K \le N, n \le N$	np, donde $p=\frac{K}{N}$	$np(1-p)\left(\frac{N-m}{N-1}\right)$	4-8
Poisson	$\frac{e^{-\lambda}\lambda^x}{x!}, x=0, 1, 2, \ldots, \lambda > 0$	λ	λ	4–9
Continua				
Uniforme	$\frac{1}{b-a}, a \le x \le b$	$\frac{(b+a)}{2}$	$\frac{(b+a)^2}{12}$	5-5
Normal	$\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$	μ	σ^2	5-6
Exponencial	$\lambda e^{-\lambda x}, x \ge 0 \quad \lambda > 0$	$1/\lambda$	$1/\lambda^2$	5-9
Erlang	$\frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!}, x > 0, r=1, 2, \ldots$	r/λ	r/λ^2	5–10.1
Gamma	$\frac{\lambda x^{r-1} e^{-\lambda x}}{\Gamma(r)}, x > 0, r > 0, \lambda > 0$	r/λ	r/λ^2	5–10.2
Weibull	$\frac{\beta}{\delta}\left(\frac{x}{\delta}\right)^{\beta-1} e^{-(x/\delta)^\beta}$, $x > 0, \beta > 0, \delta > 0$	$\delta\Gamma\left(1+\frac{1}{\beta}\right)$	$\delta^2\Gamma\left(1+\frac{2}{\beta}\right) - \delta^2\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2$	5-11

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du$$

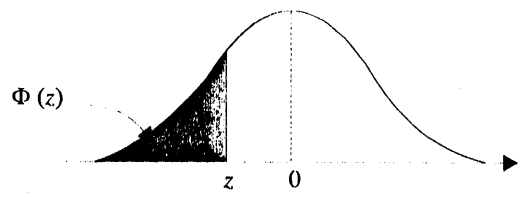


Tabla II Distribución normal estándar acumulada

z	−0.09	−0.08	−0.07	−0.06	− 0.05	−0.04	−0.03	−0.02	−0.01	−0.00
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687
−3.1	0.000711	0.000736	0.000762	0.000789	0.000816	0.000845	0.000874	0.000904	0.000935	0.000968
−3.0	0.001001	0.001035	0.001070	0.001107	0.001144	0.001183	0.001223	0.001264	0.001306	0.001350
−2.9	0.001395	0.001441	0.001489	0.001538	0.001589	0.001641	0.001695	0.001750	0.001807	0.001866
−2.8	0.001926	0.001988	0.002052	0.002118	0.002186	0.002256	0.002327	0.002401	0.002477	0.002555
−2.7	0.002635	0.002718	0.002803	0.002890	0.002980	0.003072	0.003167	0.003264	0.003364	0.003467
−2.6	0.003573	0.003681	0.003793	0.003907	0.004025	0.004145	0.004269	0.004396	0.004527	0.004661
−2.5	0.004799	0.004940	0.005085	0.005234	0.005386	0.005543	0.005703	0.005868	0.006037	0.006210
−2.4	0.006387	0.006569	0.006756	0.006947	0.007143	0.007344	0.007549	0.007760	0.007976	0.008198
−2.3	0.008424	0.008656	0.008894	0.009137	0.009387	0.009642	0.009903	0.010170	0.010444	0.010724
−2.2	0.011011	0.011304	0.011604	0.011911	0.012224	0.012545	0.012874	0.013209	0.013553	0.013903
−2.1	0.014262	0.014629	0.015003	0.015386	0.015778	0.016177	0.016586	0.017003	0.017429	0.017864
−2.0	0.018309	0.018763	0.019226	0.019699	0.020182	0.020675	0.021178	0.021692	0.022216	0.022750
−1.9	0.023295	0.023852	0.024419	0.024998	0.025588	0.026190	0.026803	0.027429	0.028067	0.028717
−1.8	0.029379	0.030054	0.030742	0.031443	0.032157	0.032884	0.033625	0.034379	0.035148	0.035930
−1.7	0.036727	0.037538	0.038364	0.039204	0.040059	0.040929	0.041815	0.042716	0.043633	0.044565
−1.6	0.045514	0.046479	0.047460	0.048457	0.049471	0.050503	0.051551	0.052616	0.053699	0.054799
−1.5	0.055917	0.057053	0.058208	0.050380	0.060571	0.061780	0.063008	0.064256	0.065522	0.066807
−1.4	0.068112	0.069437	0.070781	0.072145	0.073529	0.074934	0.076359	0.077804	0.079270	0.080757
−1.3	0.082264	0.083793	0.085343	0.086915	0.088508	0.090123	0.091759	0.093418	0.095098	0.096801
−1.2	0.098525	0.100273	0.102042	0.103835	0.105650	0.107488	0.109349	0.111233	0.113140	0.115070
−1.1	0.117023	0.119000	0.121001	0.123024	0.125072	0.127143	0.129238	0.131357	0.133500	0.135666
−1.0	0.137857	0.140071	0.142310	0.144572	0.146859	0.149170	0.151505	0.153864	0.156248	0.158655
−0.9	0.161087	0.163543	0.166023	0.168528	0.171056	0.173609	0.176185	0.178786	0.181411	0.184060
−0.8	0.186733	0.189430	0.192150	0.194894	0.197662	0.200454	0.203269	0.206108	0.208970	0.211855
−0.7	0.214764	0.217695	0.220650	0.223627	0.226627	0.229650	0.232695	0.235762	0.238852	0.241964
−0.6	0.245097	0.248252	0.251429	0.254627	0.257846	0.261086	0.264347	0.267629	0.270931	0.274253
−0.5	0.277595	0.280957	0.284339	0.287740	0.291160	0.294599	0.298056	0.301532	0.305026	0.308538
−0.4	0.312067	0.315614	0.319178	0.322758	0.326355	0.329969	0.333598	0.337243	0.340903	0.344578
−0.3	0.348268	0.351973	0.355691	0.359424	0.363169	0.366928	0.370700	0.374484	0.378281	0.382089
−0.2	0.385908	0.389739	0.393580	0.397432	0.401294	0.405165	0.409046	0.412936	0.416834	0.420740
−0.1	0.424655	0.428576	0.432505	0.436441	0.440382	0.444330	0.448283	0.452242	0.456205	0.460172
0.0	0.464144	0.468119	0.472097	0.476078	0.480061	0.484047	0.488033	0.492022	0.496011	0.500000

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} \, du$$

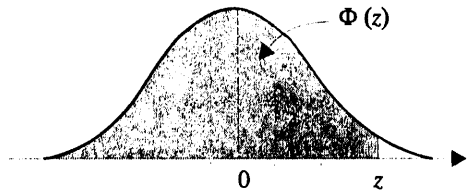


Tabla II Distribución normal estándar acumulada (*continuación*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959071	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965621	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967

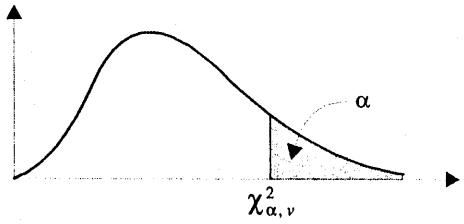


Tabla III Puntos porcentuales $\chi^2_{\alpha,\nu}$ de la distribución ji-cuadrada.

ν \ α	–995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.00+	.00+	.00+	.00+	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09-	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

ν = grados de libertad.

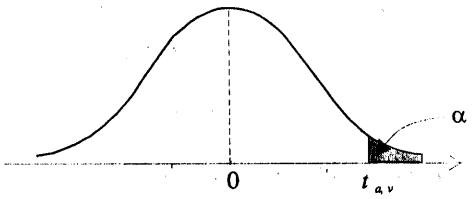
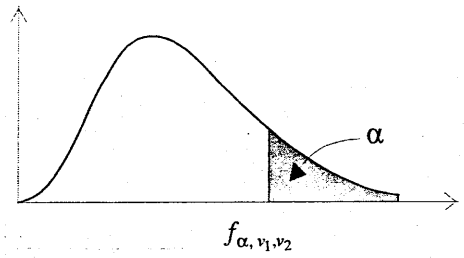


Tabla IV Puntos porcentuales $t_{\alpha,\nu}$, de la distribución t

ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	7.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = grados de libertad.



$f_{0.25, v_1, v_2}$

Tabla V Puntos porcentuales f_{α}, v_1, v_2, de la distribución F

v_2 \ v_1	Grados de libertad del numerador (v_1)																		
Grados de libertad del denominador (v_2)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.41	9.49	9.58	9.63	9.67	9.71	9.76	9.80	9.85
2	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.41	3.43	3.43	3.44	3.45	3.46	3.47	3.48
3	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.46	2.46	2.46	2.47	2.47	2.47	2.47	2.47
4	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
5	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.88	1.88	1.88	1.88	1.87	1.87	1.87
6	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.76	1.76	1.75	1.75	1.75	1.74	1.74	1.74
7	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.70	1.69	1.68	1.68	1.67	1.67	1.66	1.66	1.65	1.65	1.65
8	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.62	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58
9	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.54	1.53	1.53
10	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.53	1.52	1.52	1.51	1.51	1.50	1.49	1.48
11	1.47	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.47	1.46	1.45
12	1.46	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.45	1.44	1.43	1.42
13	-1.45	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.42	1.41	1.40
14	1.44	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.41	1.41	1.40	1.39	1.38
15	1.43	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.46	1.45	1.44	1.43	1.41	1.41	1.40	1.39	1.38	1.37	1.36
16	1.42	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.44	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
17	1.42	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33
18	1.41	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.34	133	1.32
19	1.41	1.49	1.49	1.47	1.46	1.44	1.43	1.42	1.41	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.33	1.32	1.30
20	1.40	1.49	1.48	1.47	1.45	1.44	1.43	1.42	1.41	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31	1.29
21	1.40	1.48	1.48	1.46	1.44	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28
22	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28
23	1.39	1.47	1.47	1.45	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28	1.27
24	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.36	1.35	1.33	1.32	1.31	1.30	1.29	1.28	1.26
25	1.39	1.47	1.46	1.44	1.42	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.29	1.28	1.27	1.25
26	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.35	1.34	1.32	1.31	1.30	1.29	1.28	1.26	1.25
27	1.38	1.46	1.45	1.43	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.33	1.32	1.31	1.30	1.28	1.27	1.26	1.24
28	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.31	1.30	1.29	1.28	1.27	1.25	1.24
29	1.38	1.45	1.45	1.43	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.32	1.31	1.30	1.29	1.27	1.26	1.25	1.23
30	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.34	1.32	1.30	1.29	1.28	1.27	1.26	1.24	1.23
40	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.31	1.30	1.28	1.26	1.25	1.24	1.22	1.21	1.19
60	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.27	1.25	1.24	1.22	1.21	1.19	1.17	1.15

Tabla V Puntos porcentuales de la distribución F (*continuación*)

$f_{0.10,\nu_1,\nu_2}$

ν_2 \ ν_1	Grados de libertad del numerador (ν_1)																		
Grados de libertad del denominador (ν_2)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.2	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.03	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

Tabla V Puntos porcentuales de la distribución F (*continuación*)

$f_{0.05,\,\nu_1,\nu_2}$

Grados de libertad del denominador (ν_2) \ Grados de libertad del numerador (ν_1)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	4.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.94	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.55	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Tabla V Puntos porcentuales de la distribución F (*continuación*)

$f_{0.10,\,\nu_1,\nu_2}$

ν_2 \ ν_1	Grados de libertad del numerador (ν_1)																		
Grados de libertad del denominador (ν_2)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	*3.66*	*3.50*	*3.38*	*3.29*	*3.21*	*3.15*	*3.05*	*2.95*	*2.84*	2.79	2.73	2.67	*2.61*	*2.55*	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.(>6	2.96	2.86	2.76	2.70	2.(>4	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	*5.83*	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

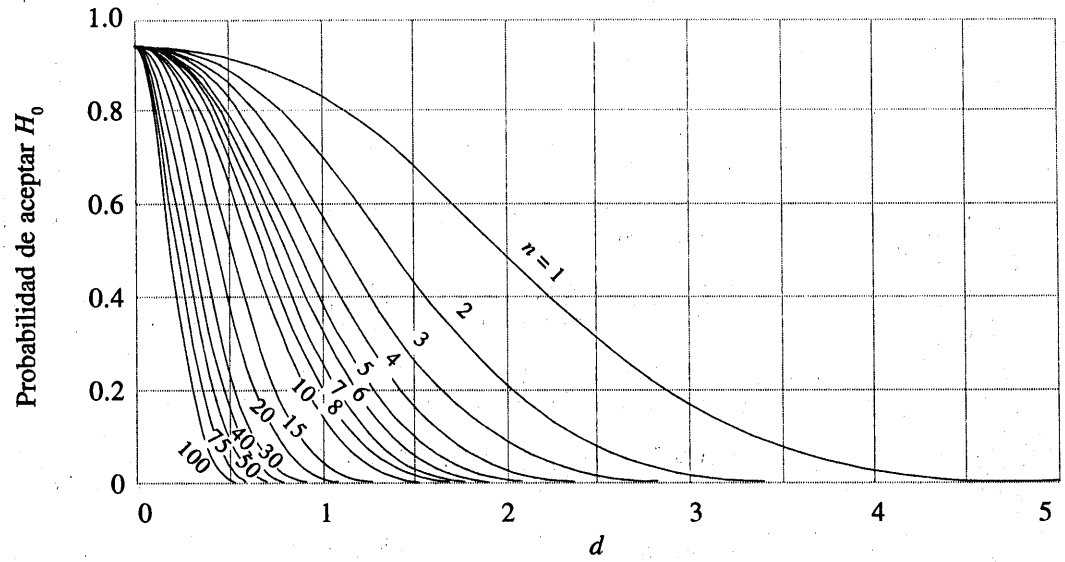
Tabla V Puntos porcentuales de la distribución F (*continuación*)

$f_{0.01,\,\nu_1,\nu_2}$

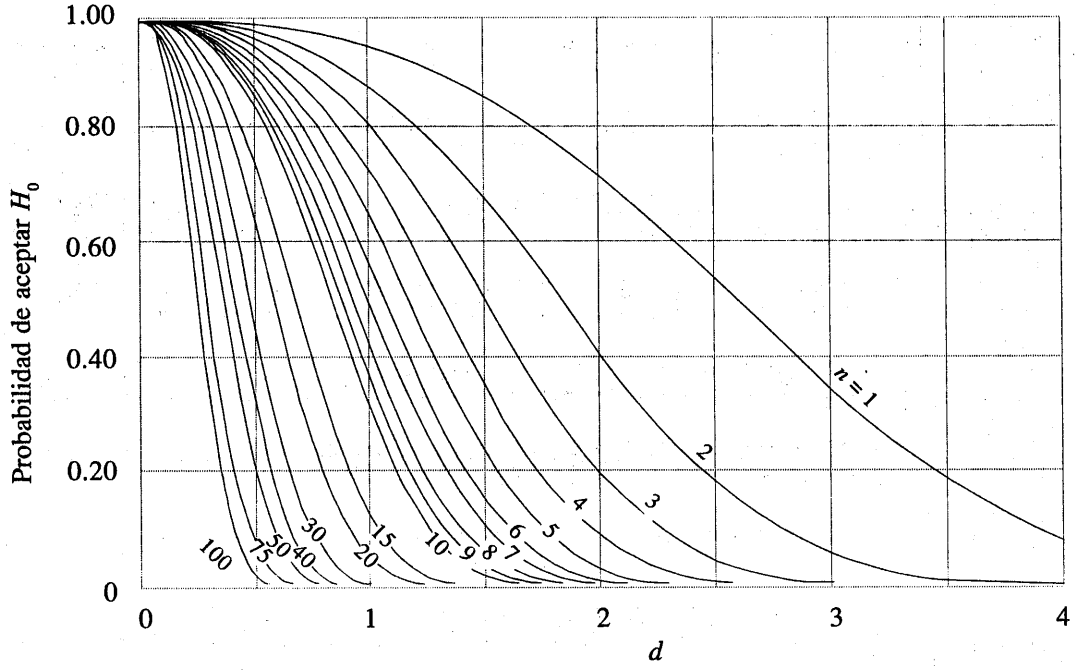
Grados de libertad del denominador (ν_2) \ Grados de libertad del numerador (ν_1)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.00	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.46
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.36	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.59
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

ν = grados de libertad.

Carta VI Curvas de operación característica



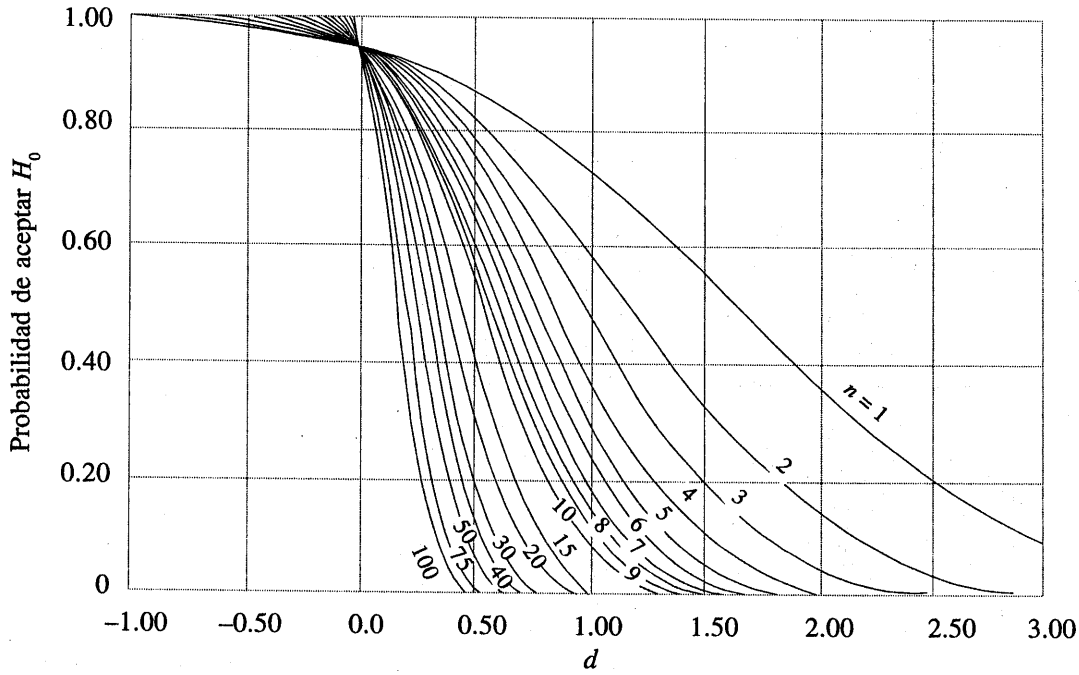
a) Curvas *OC* para diferentes valores de *n* para la prueba normal de dos colas con un nivel de significación $\alpha = 0.05$.



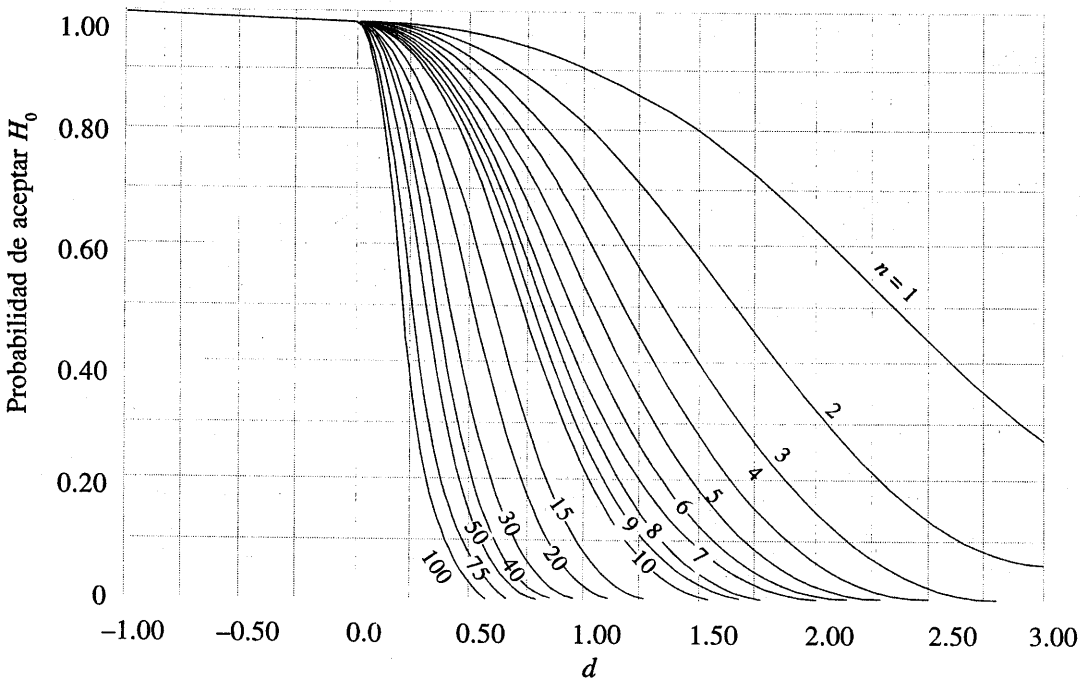
b) Curvas *OC* para diferentes valores de *n* para la prueba normal de dos colas con un nivel de significación $\alpha = 0.01$.

Fuente: Las cartas VI*a*, *e*, *f*, *k*, *m* y *q* se reproducen con permiso de "Operating Characteristics for the Common Statistical Tests of Significance", de C. L. Ferris, F. E. Grubbs y C. L. Weaver, *Annals of Mathematical Statistics*. Las cartas VI*b*, *c*, *d*, *g*, *h*, *i*, *j*, *l*, *n*, *o*, *p* y *r* se reproducen con permiso de *Engineering Statistics*, 2a. edición, de A. H. Bowker y G. J. Lieberman, Prentice-Hall.

Carta VI Curvas de operación característica (*continuación*)

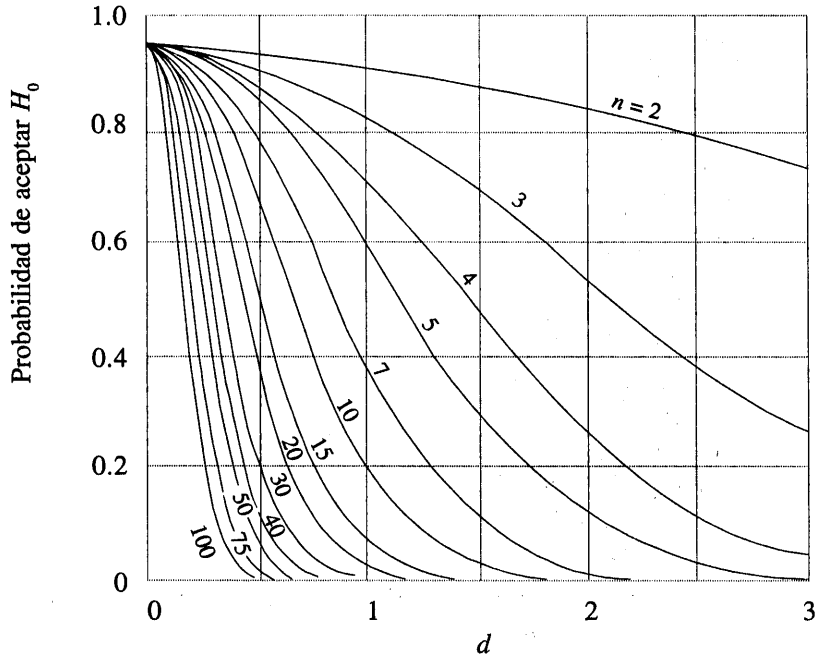


c) Curvas *OC* para diferentes valores de *n* para la prueba normal de una cola con un nivel de significación $\alpha = 0.05$.

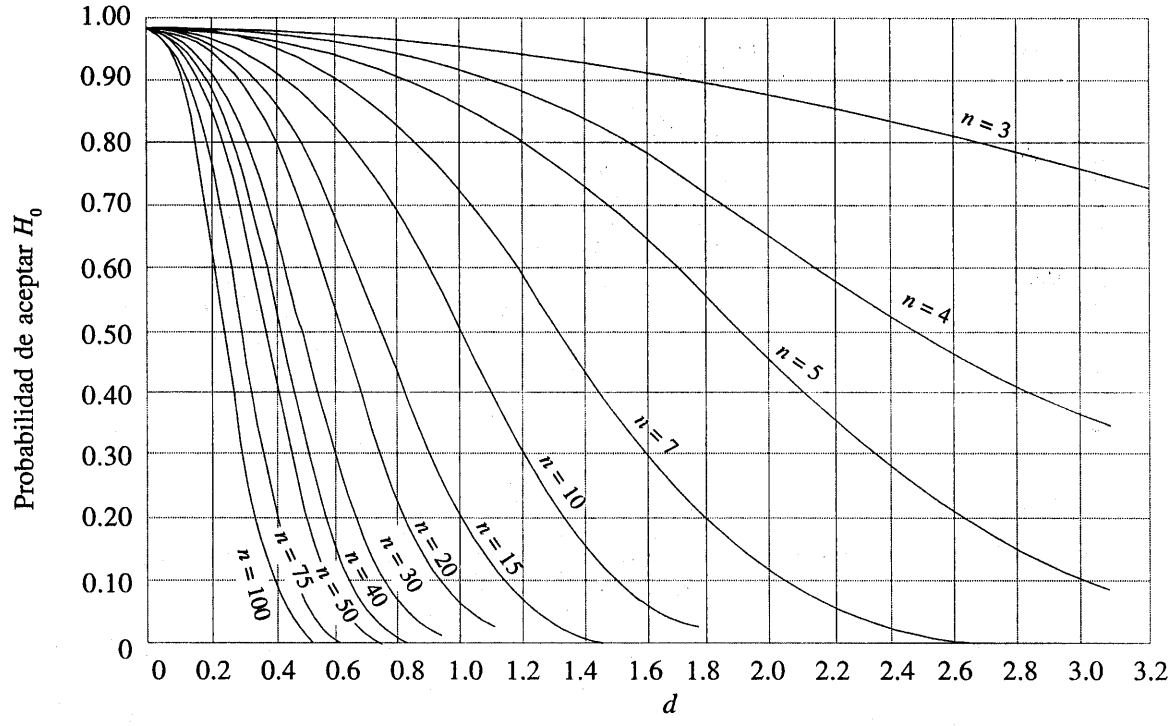


d) Curvas *OC* para diferentes valores de *n* para la prueba *t* de dos colas con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

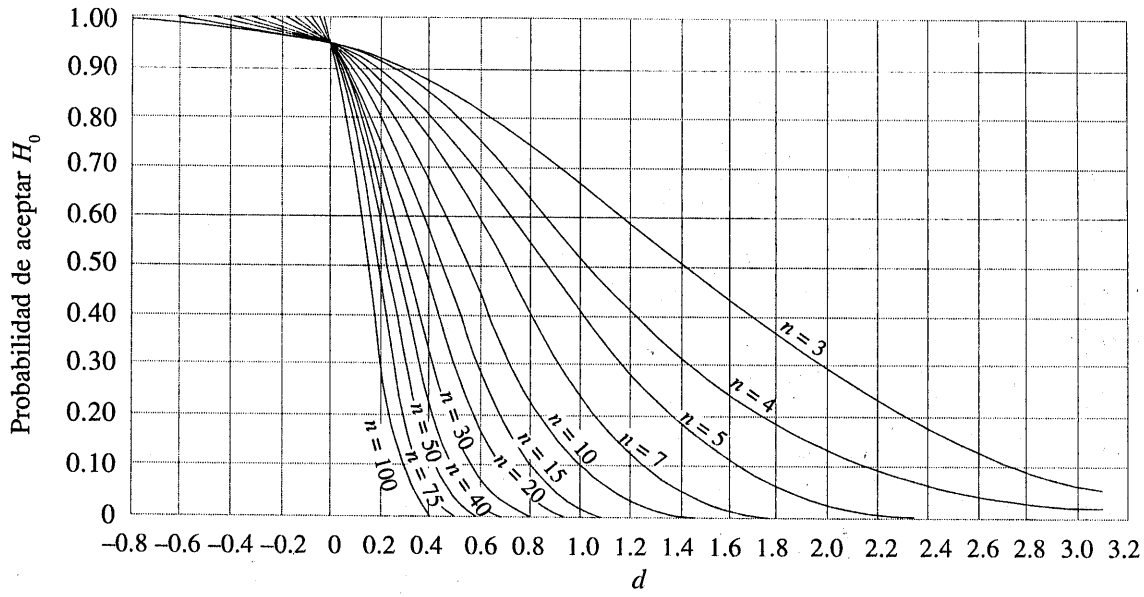


e) Curvas *OC* para diferentes valores de *n* para la prueba *t* de dos colas con un nivel de significación $\alpha = 0.05$.

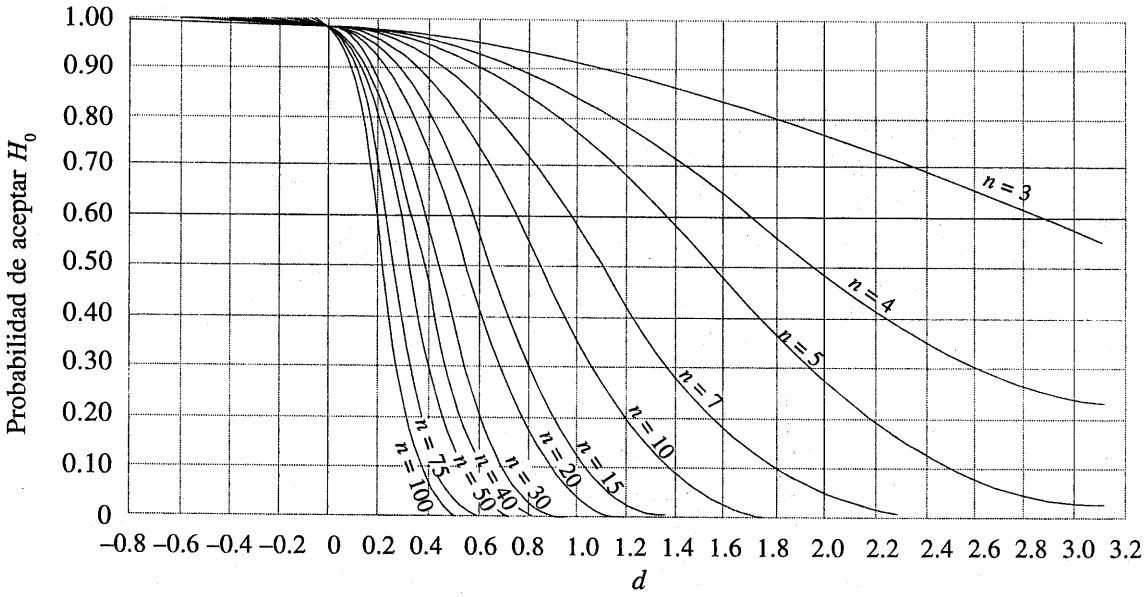


f) Curvas *OC* para diferentes valores de *n* para la prueba *t* de dos colas con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

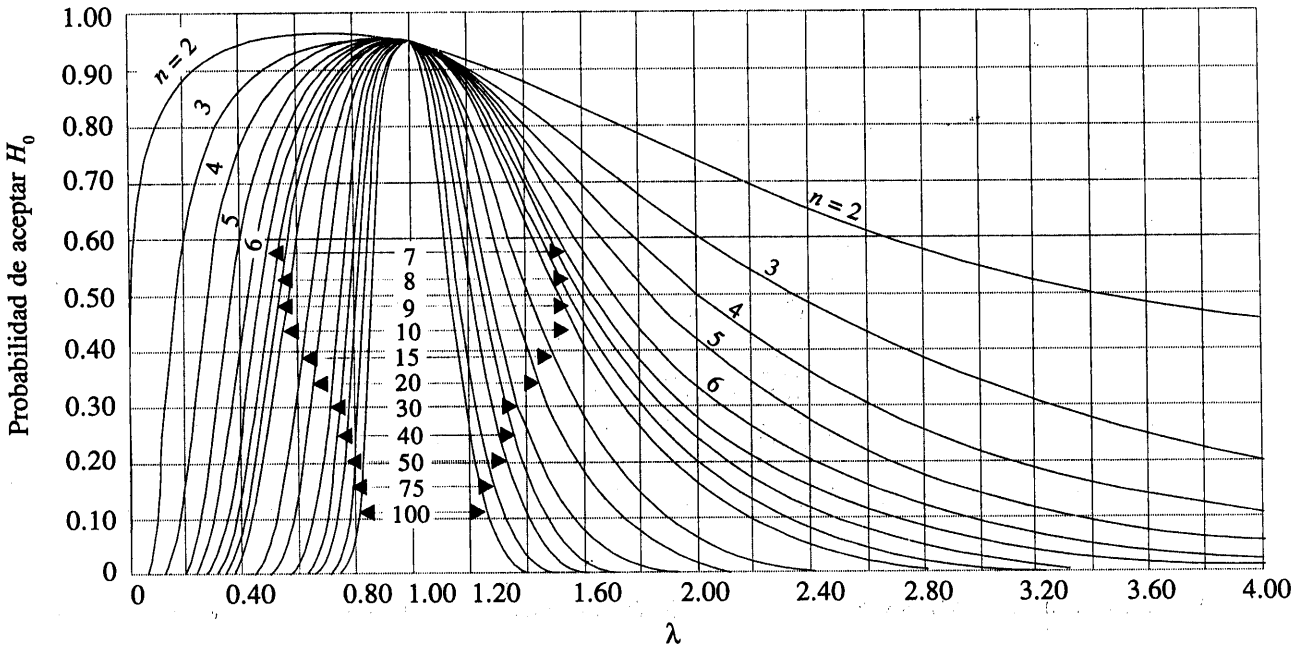


g) Curvas *OC* para diferentes valores de *n* para la prueba *t* de una cola con un nivel de significación $\alpha = 0.05$.

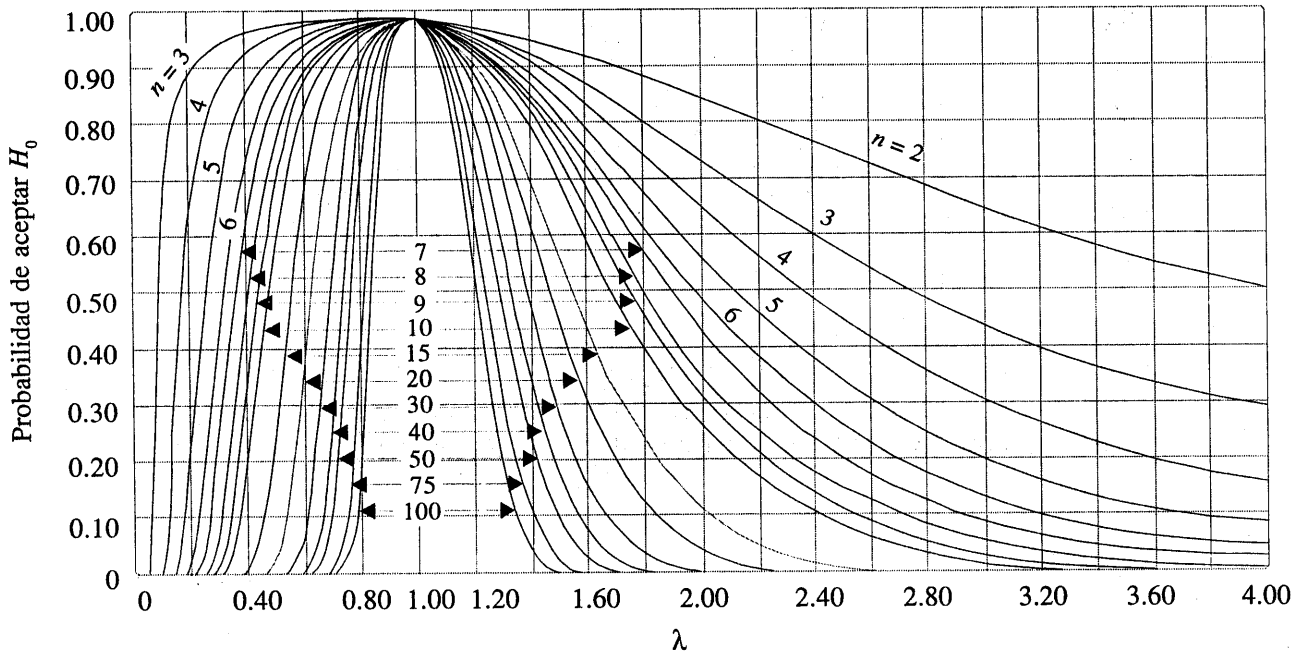


h) Curvas *OC* para diferentes valores de *n* para la prueba *t* de una cola con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

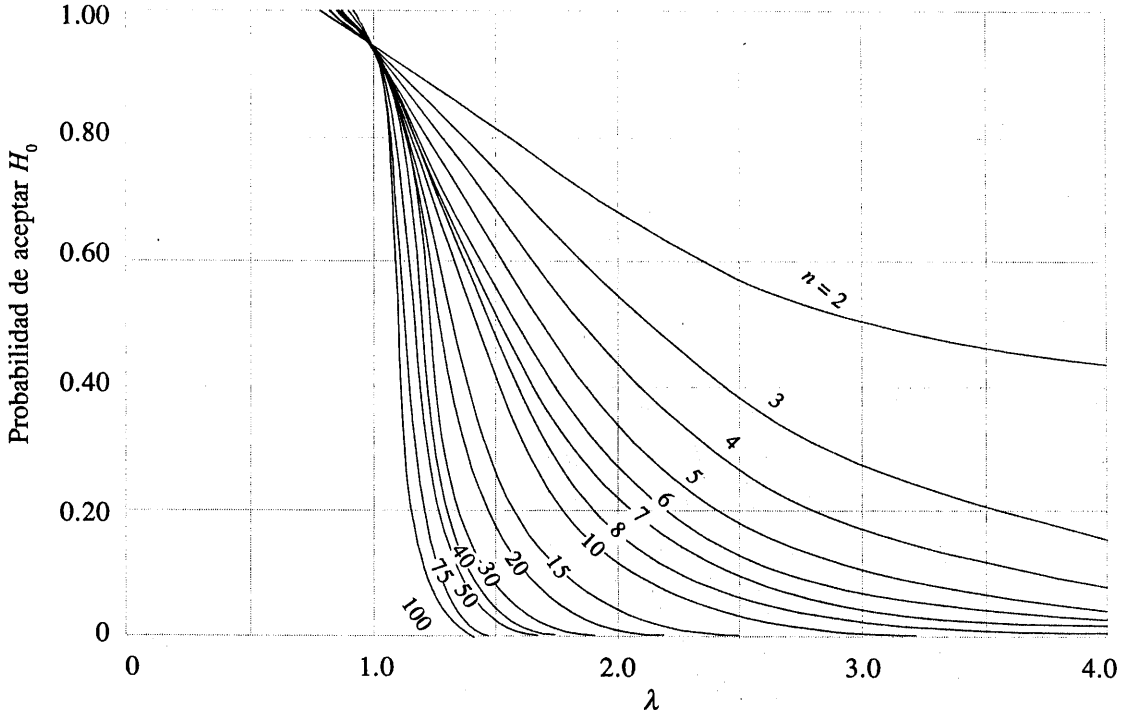


i) Curvas *OC* para diferentes valores de *n* para la prueba ji-cuadrada de dos colas con un nivel de significación $\alpha = 0.05$.

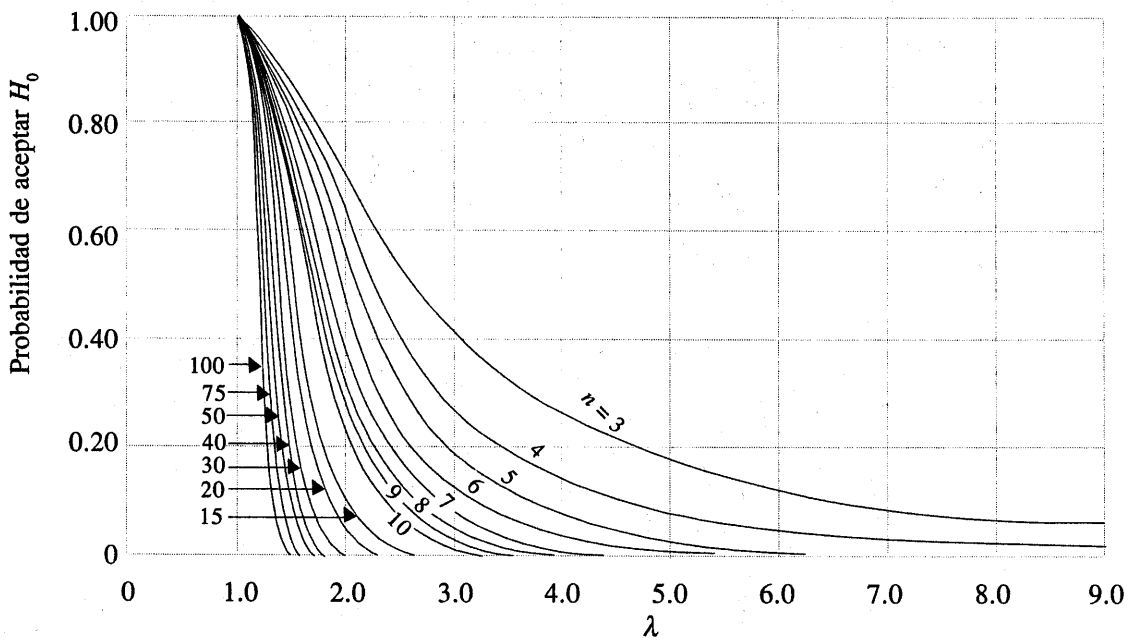


j) Curvas *OC* para diferentes valores de *n* para la prueba ji-cuadrada de dos colas con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

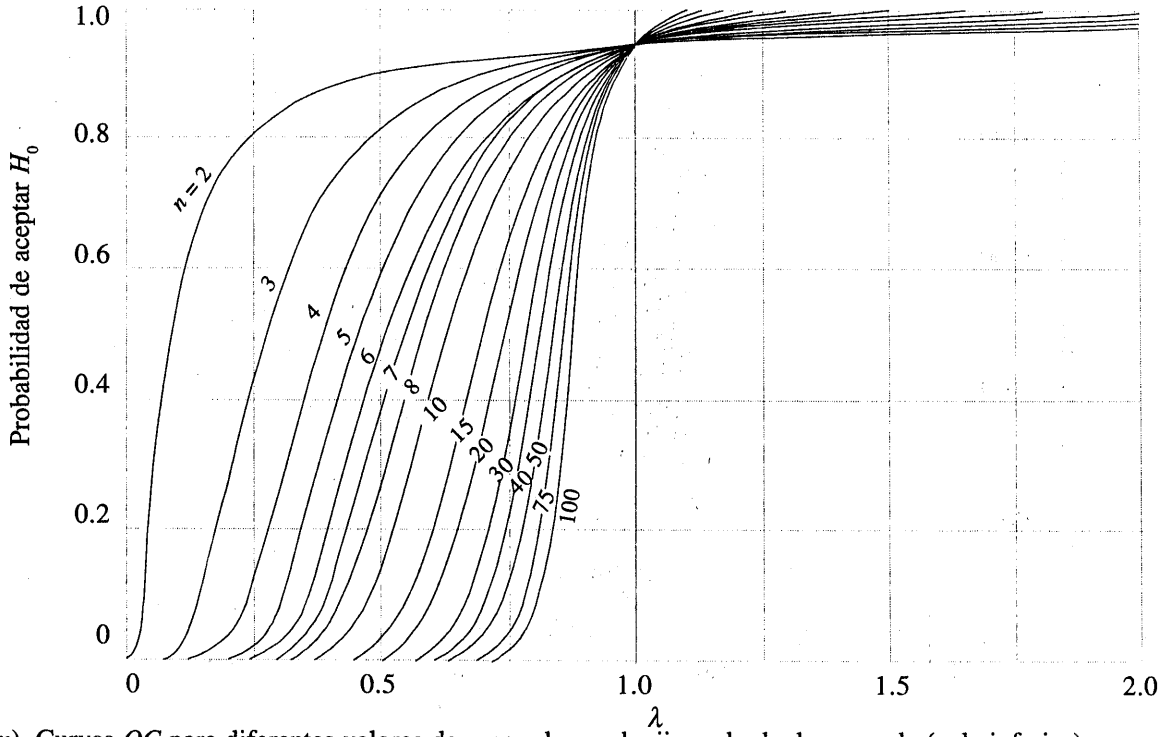


k) Curvas *OC* para diferentes valores de *n* para la prueba ji-cuadrada de una cola (cola superior) con un nivel de significación $\alpha = 0.05$.

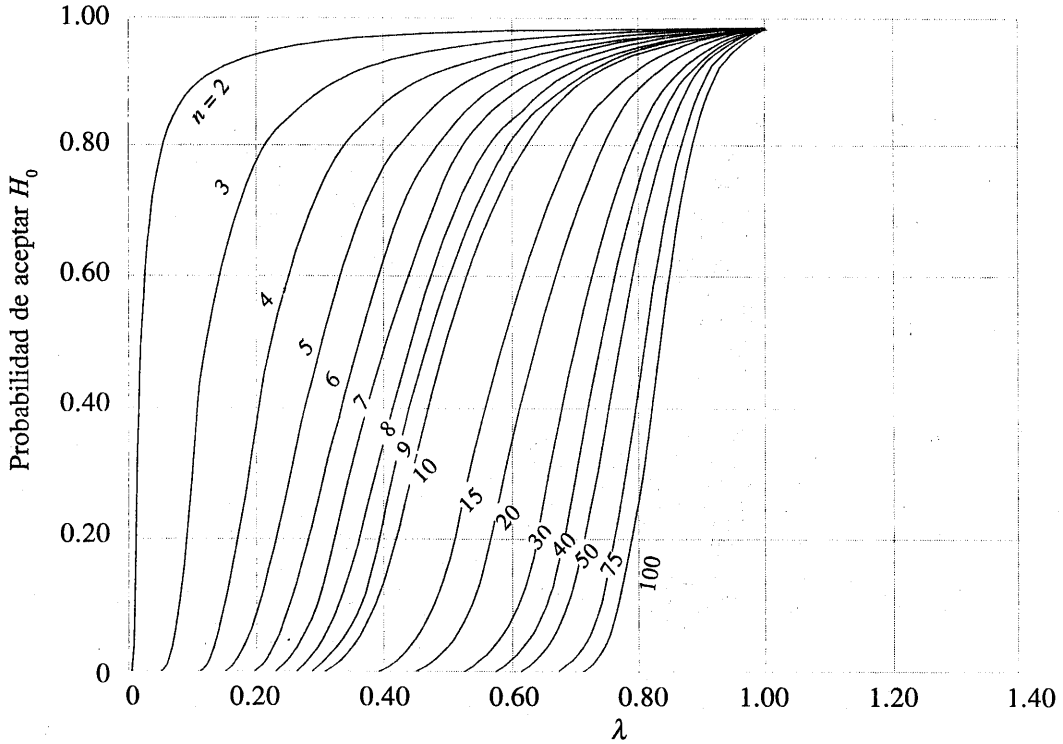


l) Curvas *OC* para diferentes valores de *n* para la prueba ji-cuadrada de una cola (cola superior) con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

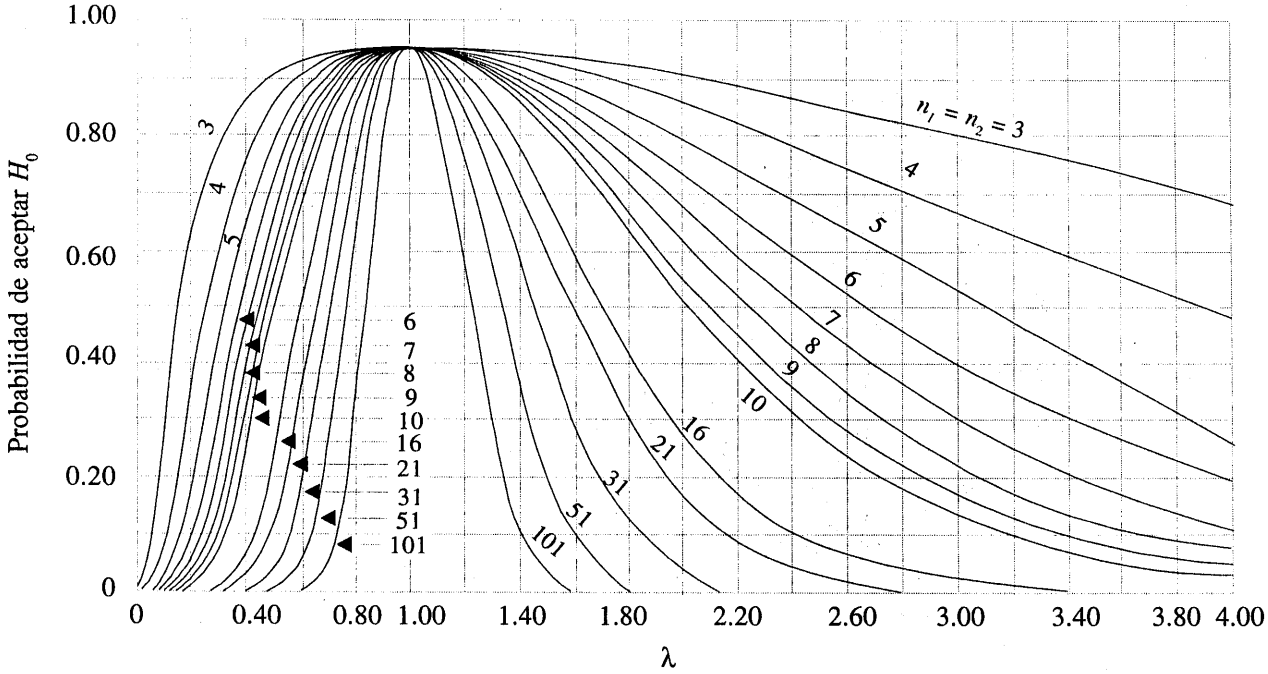


m) Curvas *OC* para diferentes valores de *n* para la prueba ji-cuadrada de una cola (cola inferior) con un nivel de significación $\alpha = 0.05$.

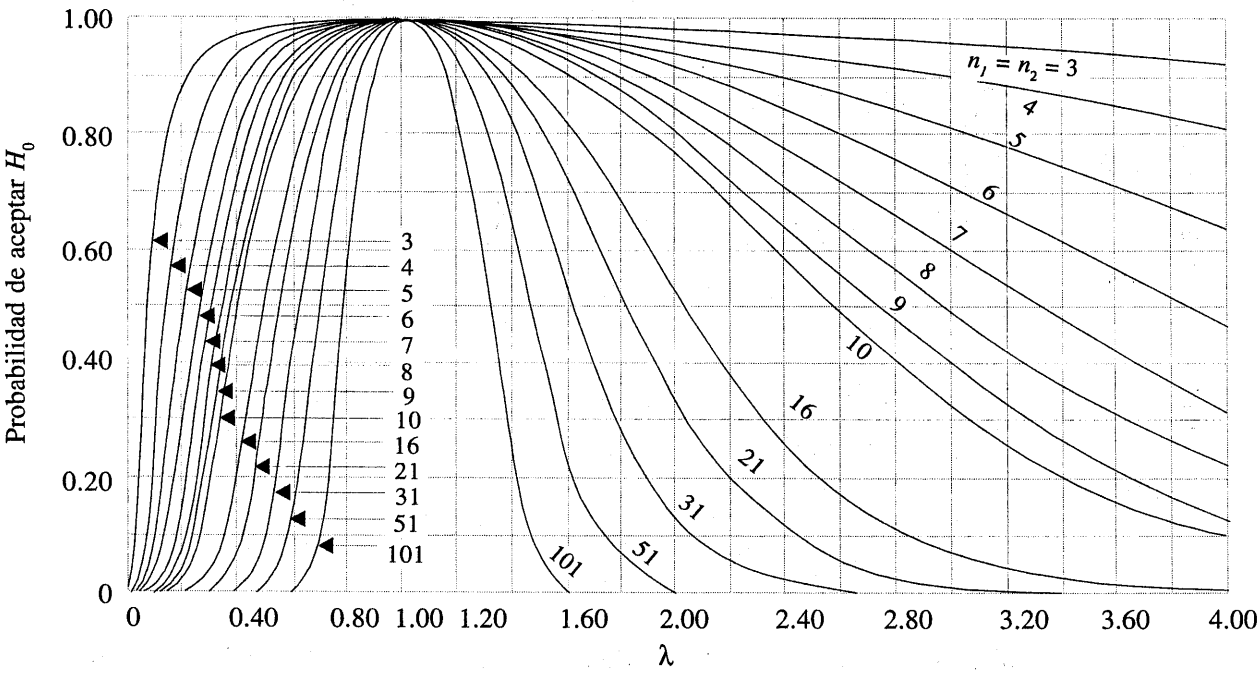


n) Curvas *OC* para diferentes valores de *n* para la prueba ji-cuadrada de una cola (cola inferior) con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

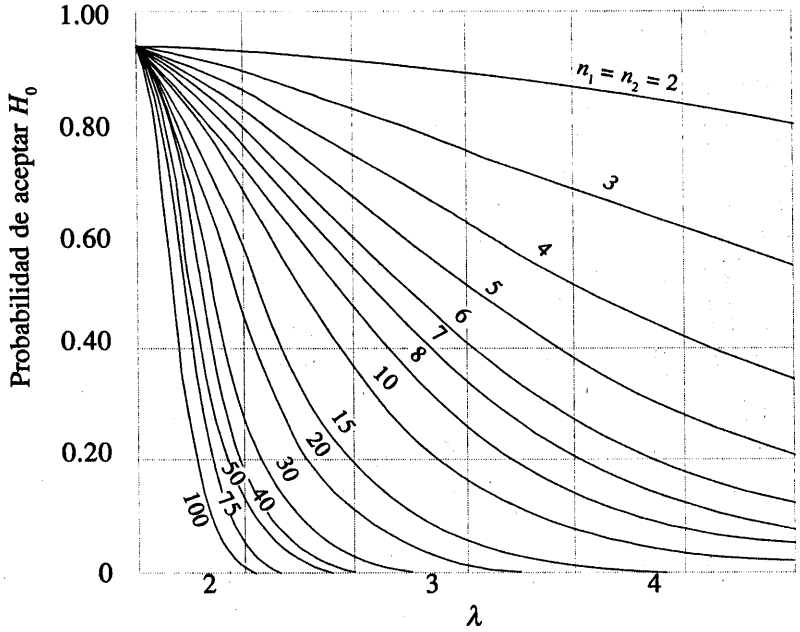


o) Curvas *OC* para diferentes valores de *n* para la prueba *F* de dos colas con un nivel de significación $\alpha = 0.05$.

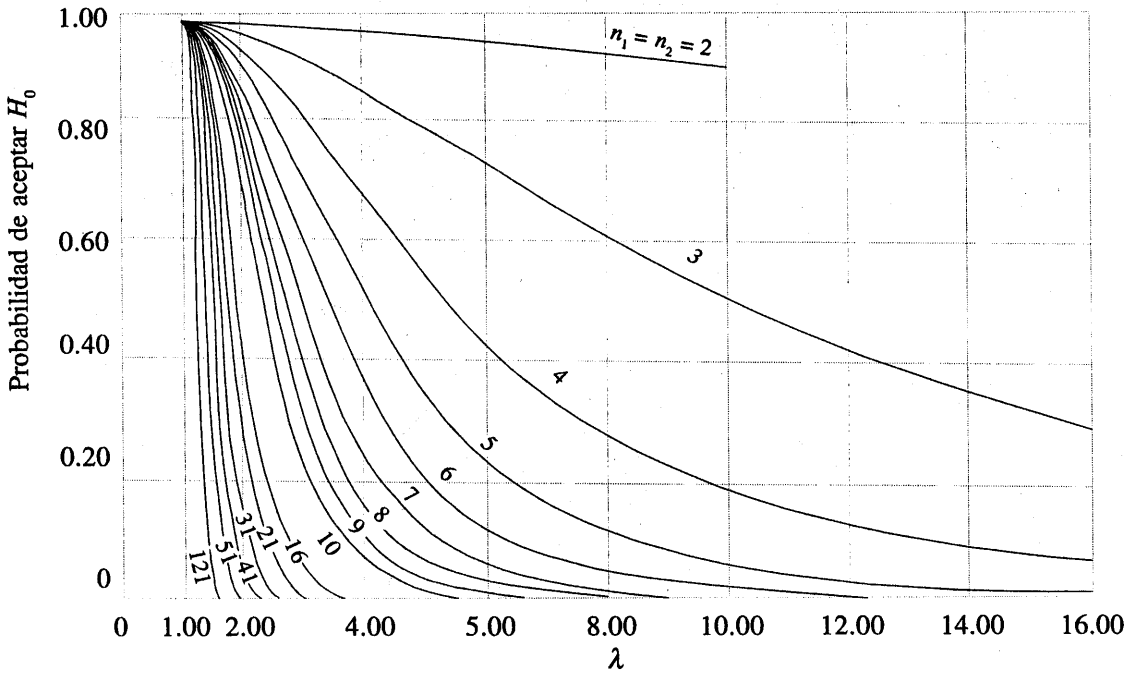


p) Curvas *OC* para diferentes valores de *n* para la prueba *F* de dos colas con un nivel de significación $\alpha = 0.01$.

Carta VI Curvas de operación característica (*continuación*)

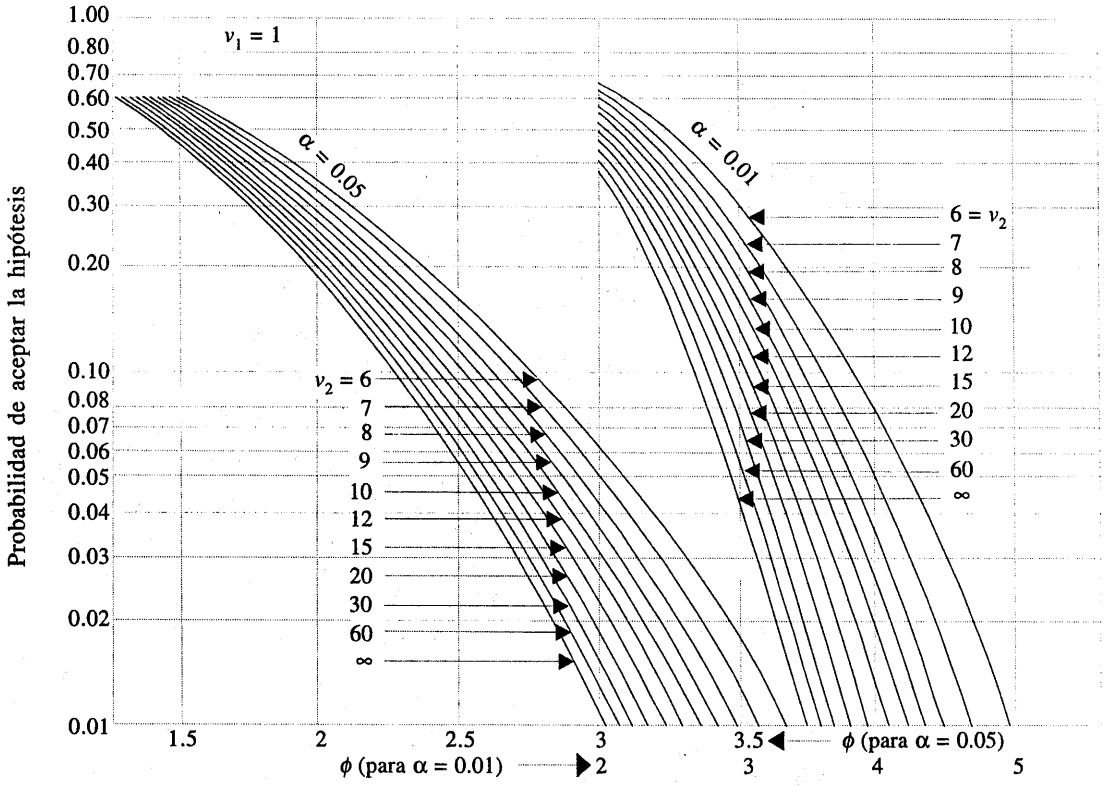


q) Curvas *OC* para diferentes valores de *n* para la prueba *F* de una cola con un nivel de significación $\alpha = 0.05$.

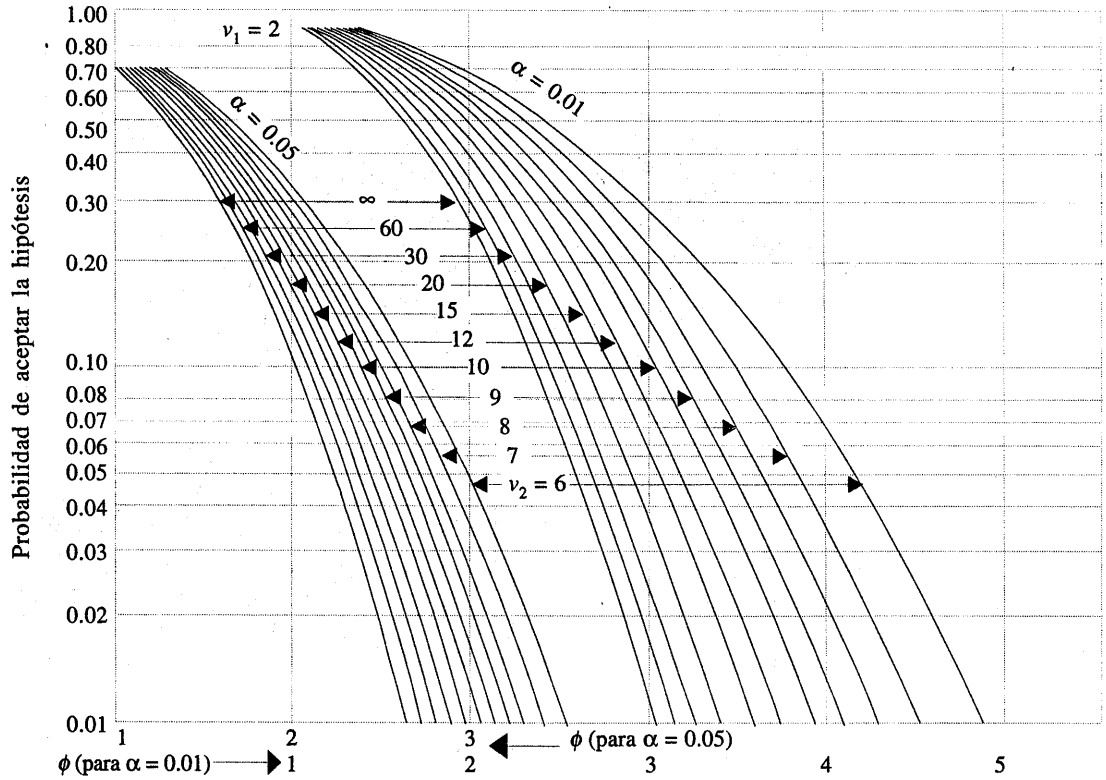


r) Curvas *OC* para diferentes valores de *n* para la prueba *F* de una cola con un nivel de significación $\alpha = 0.01$.

Carta VII Curvas de operación característica para el análisis de varianza del modelo con efectos fijos

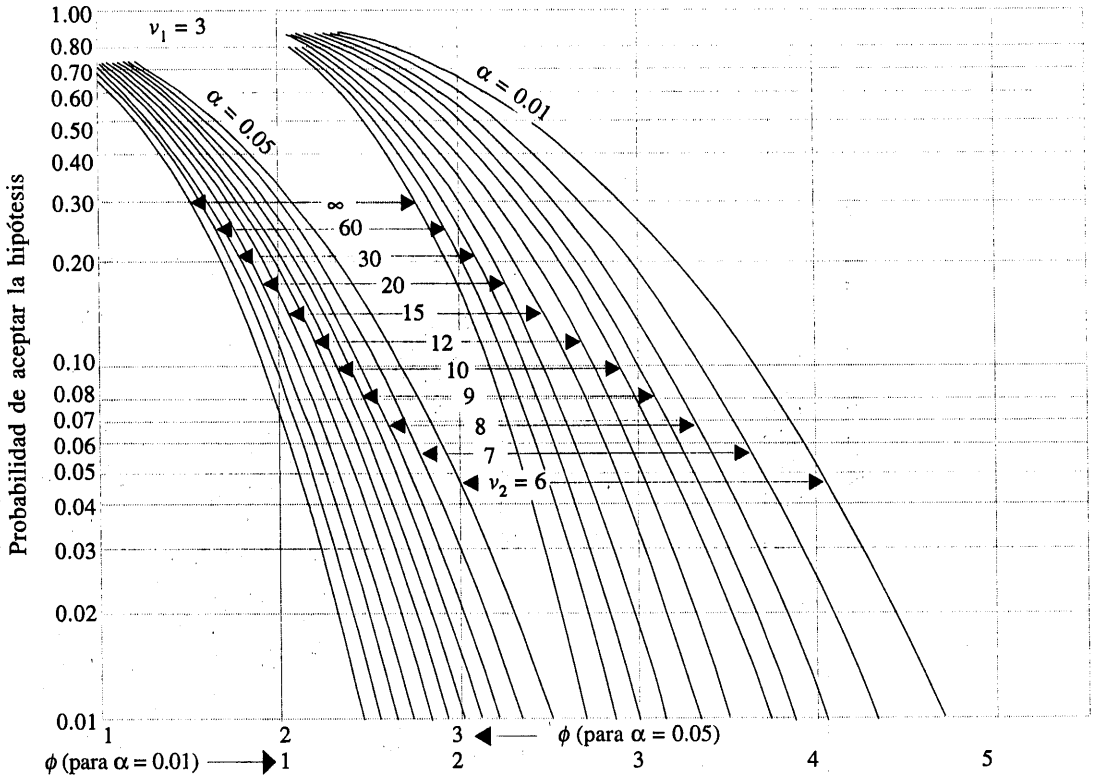


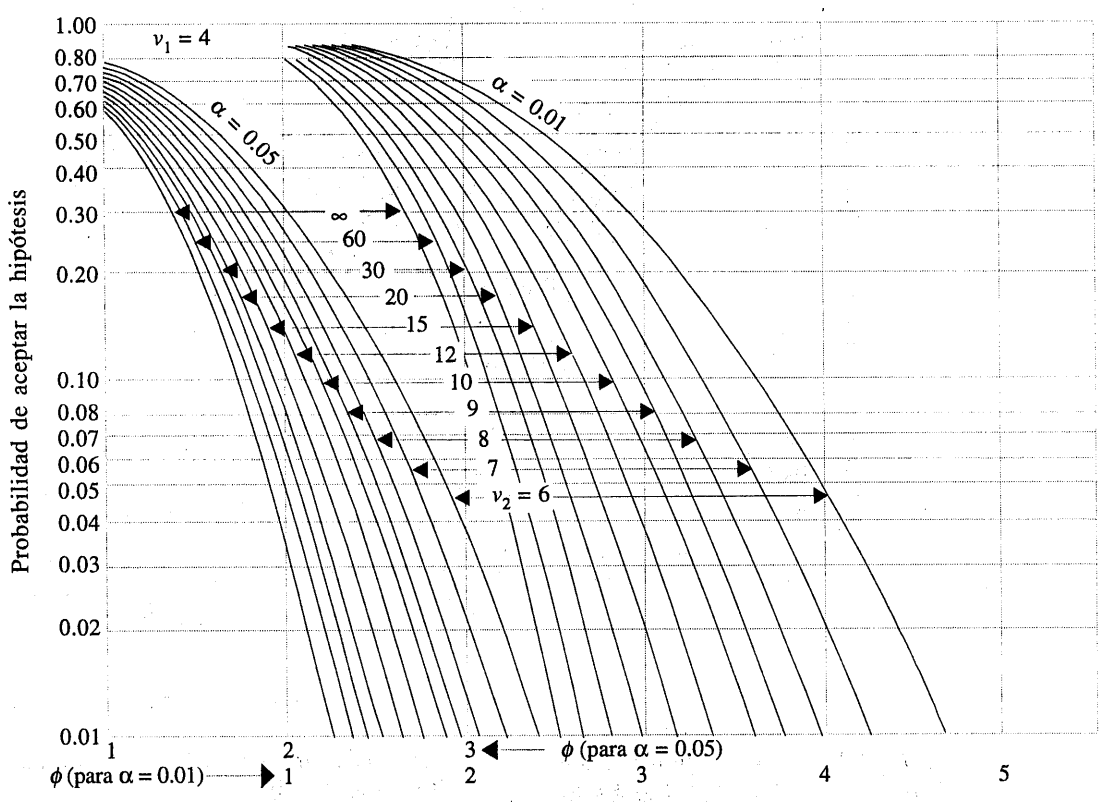
ν_1 = grados de libertad del numerador. ν_2 = grados de libertad del denominador.



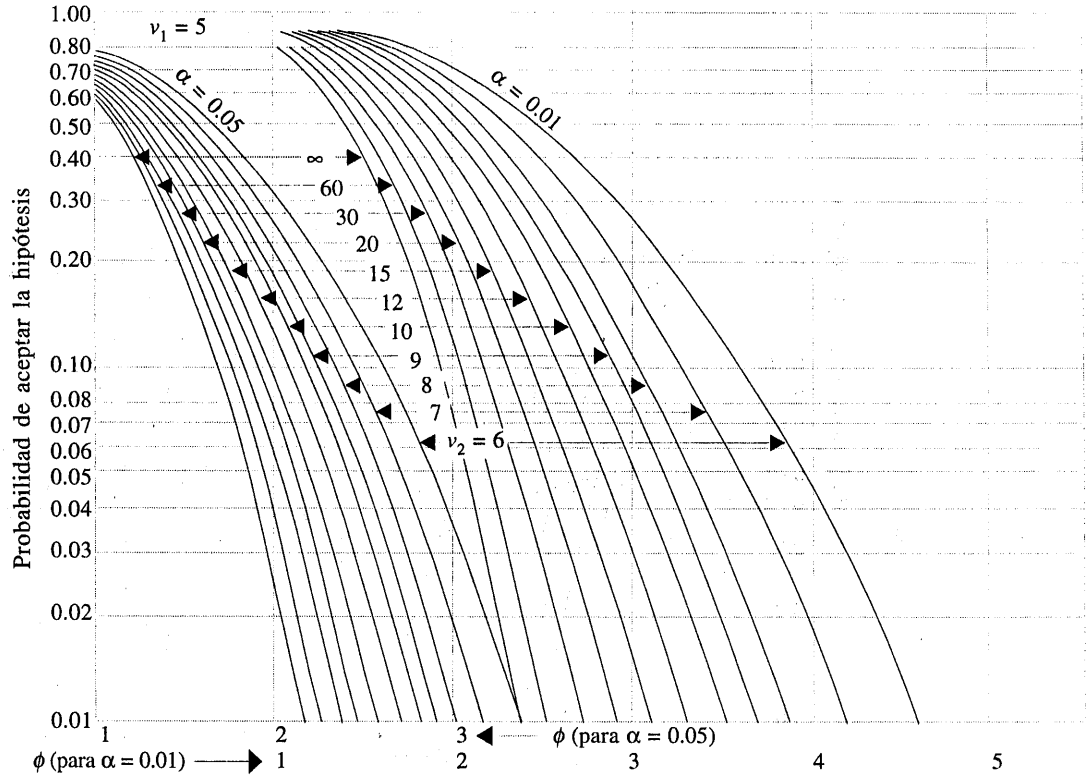
Fuente: La carta VII es una adaptación con permiso de *Biometrika Tables for Statisticians*, Vol. 2, de E. S. Pearson y H. O. Hartley, Cambridge University Press, Cambridge.

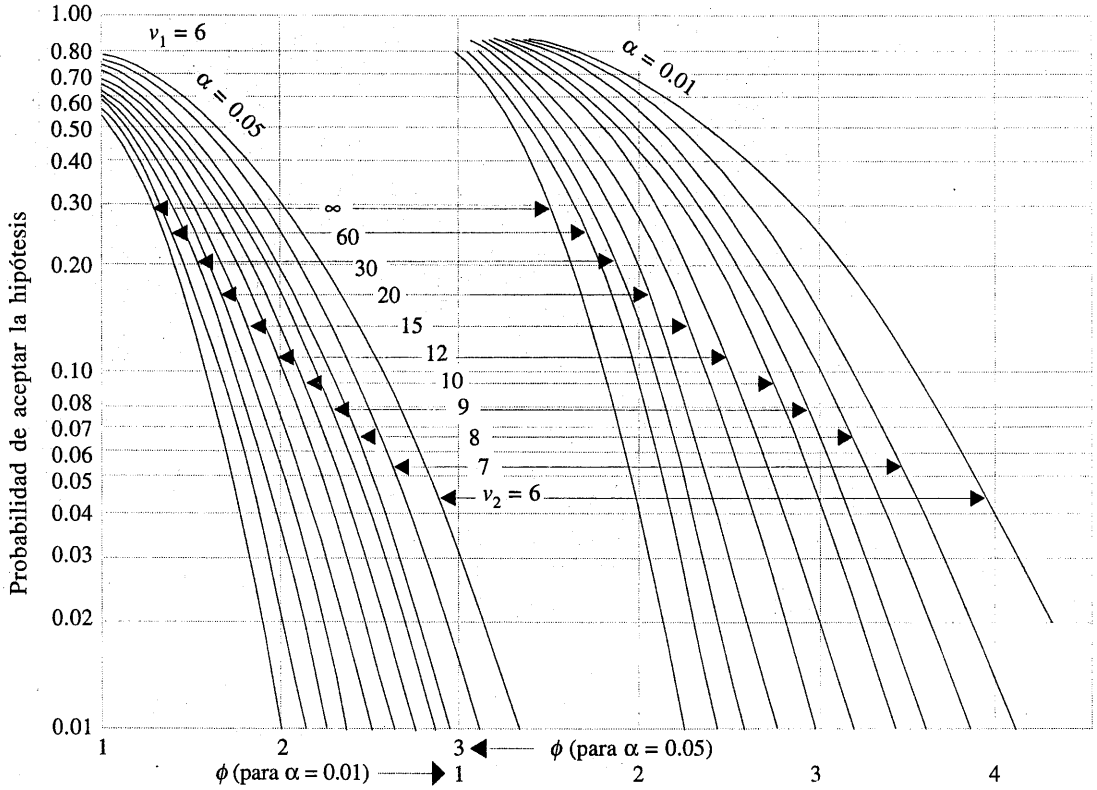
Carta VII Curvas de operación característica para el análisis de varianza del modelo con efectos fijos (*continuación*)



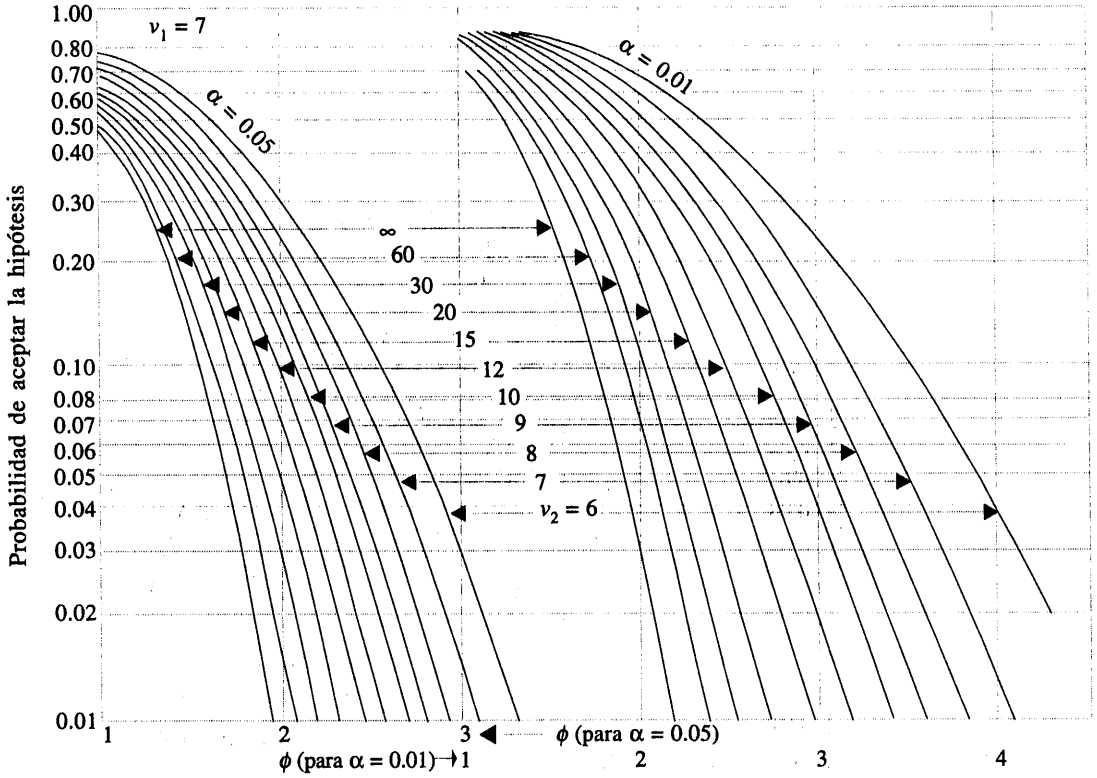


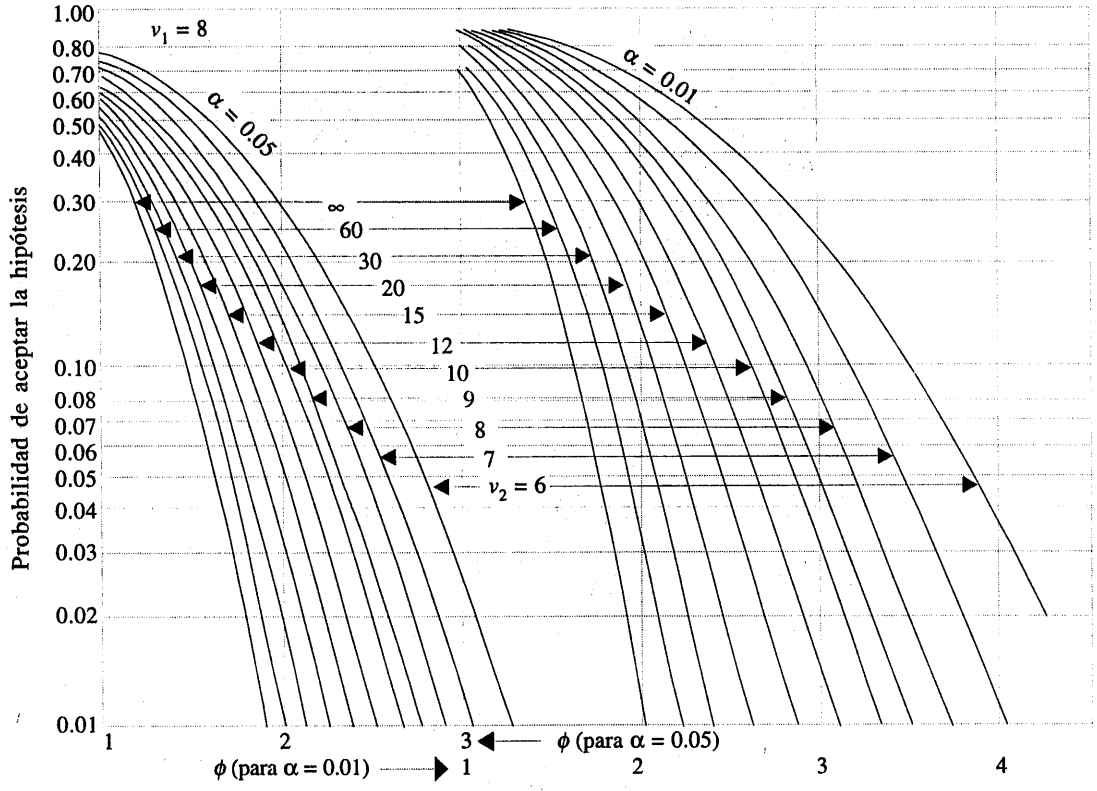
Carta VII Curvas de operación característica para el análisis de varianza del modelo con efectos fijos (*continuación*)



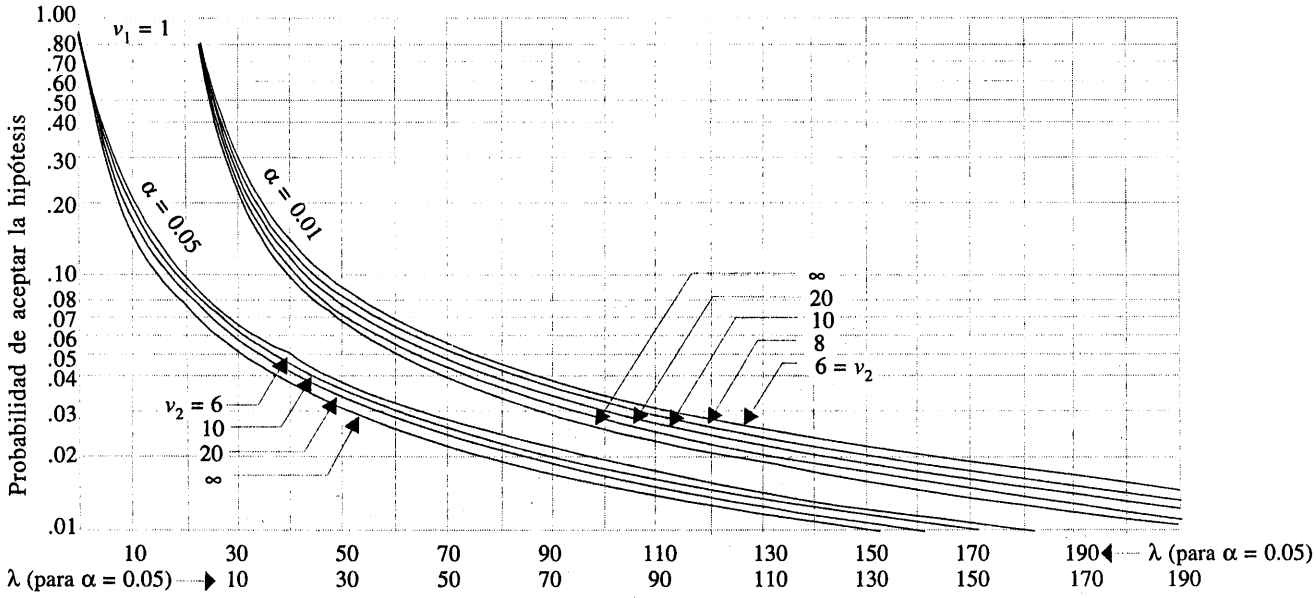


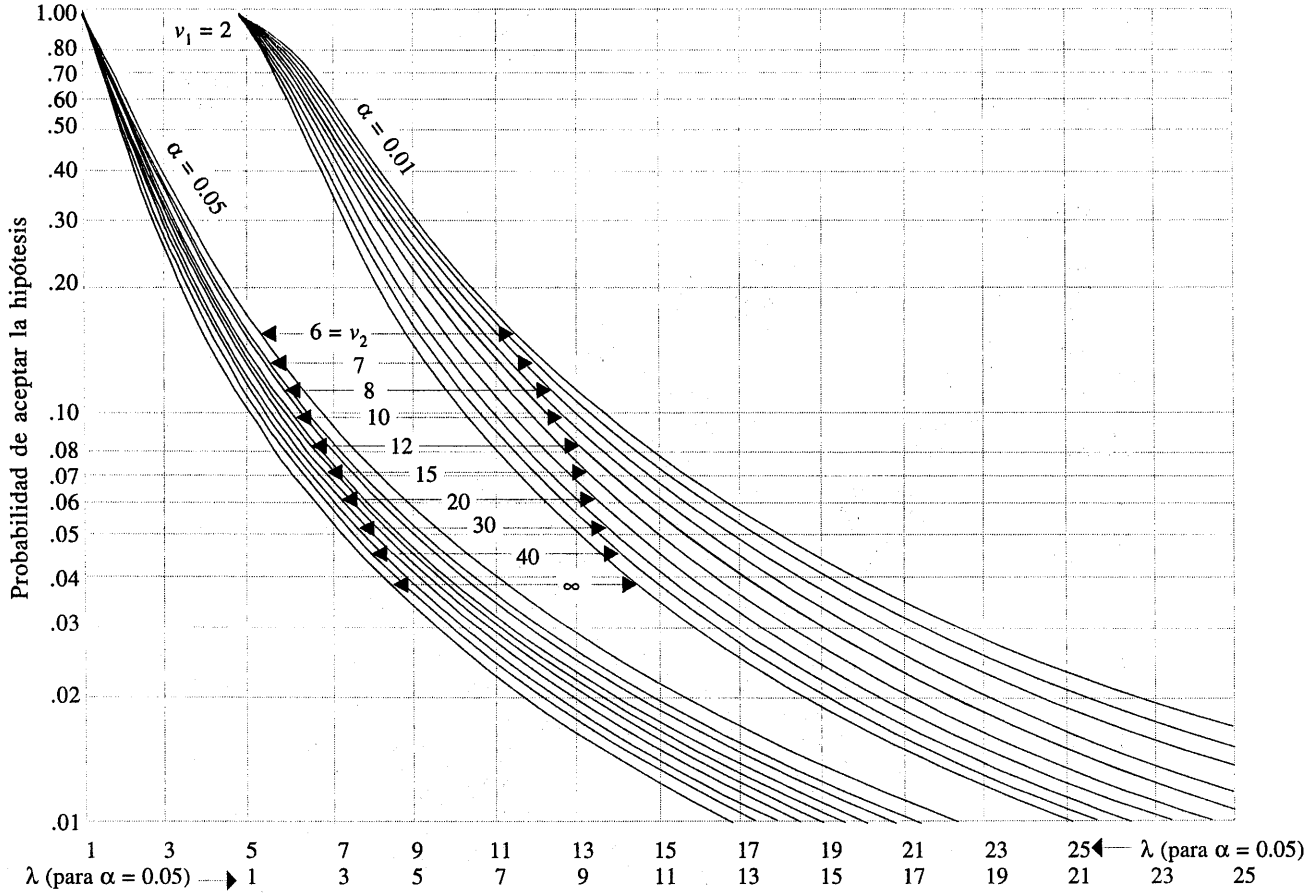
Carta VII Curvas de operación característica para el análisis de varianza del modelo con efectos fijos (*continuación*)





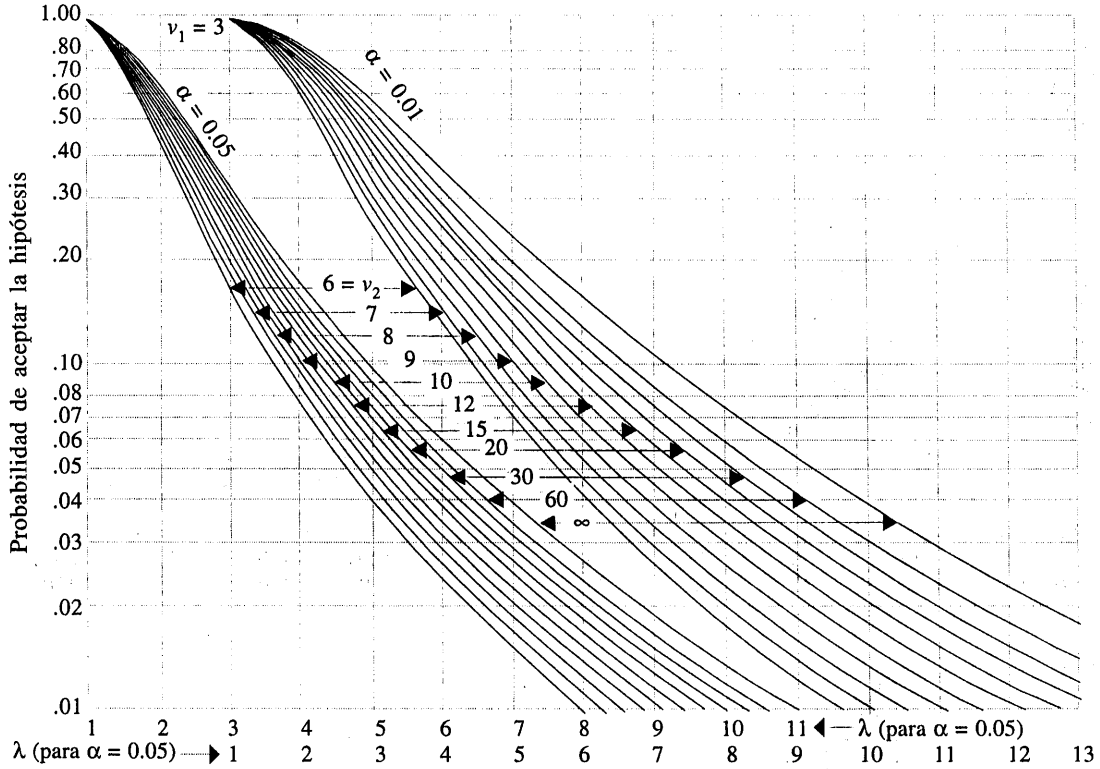
Carta VIII Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios

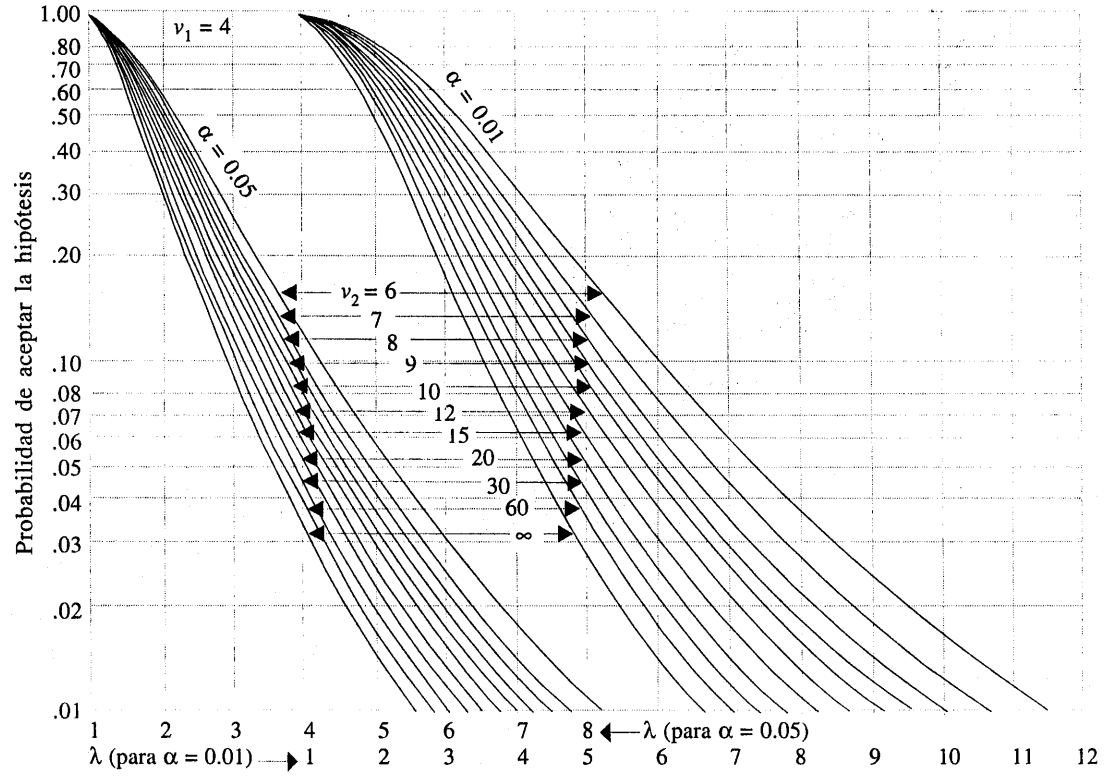




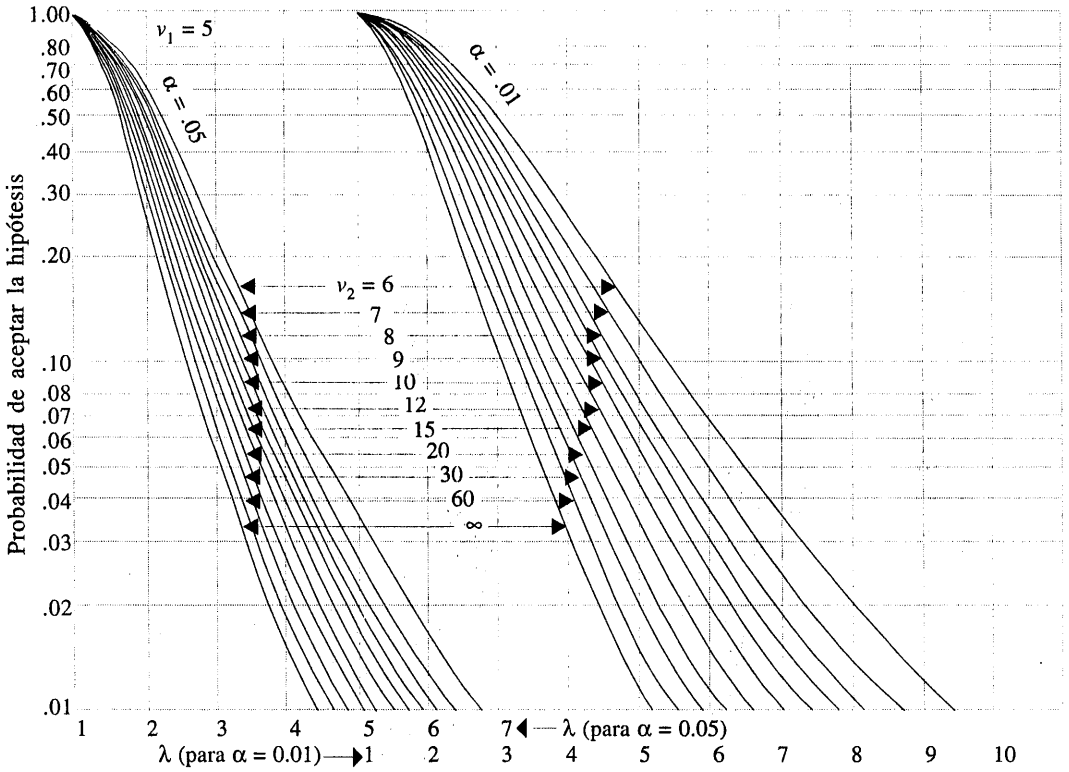
Fuente: Reproducida con permiso de *Engineering Statistics*, 2a. edición, de A. H. Bowker y G. J. Lieberman, Prentice-Hall, Englewood Cliffs, NJ.

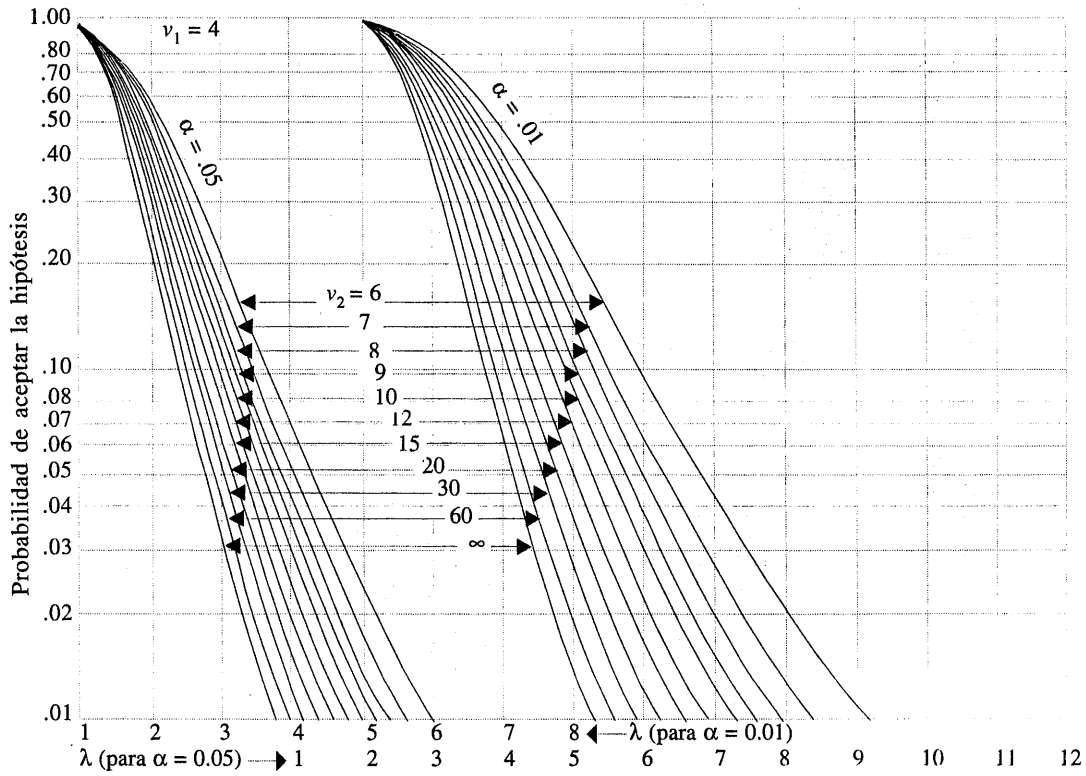
Carta VIII Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios (*continuación*)



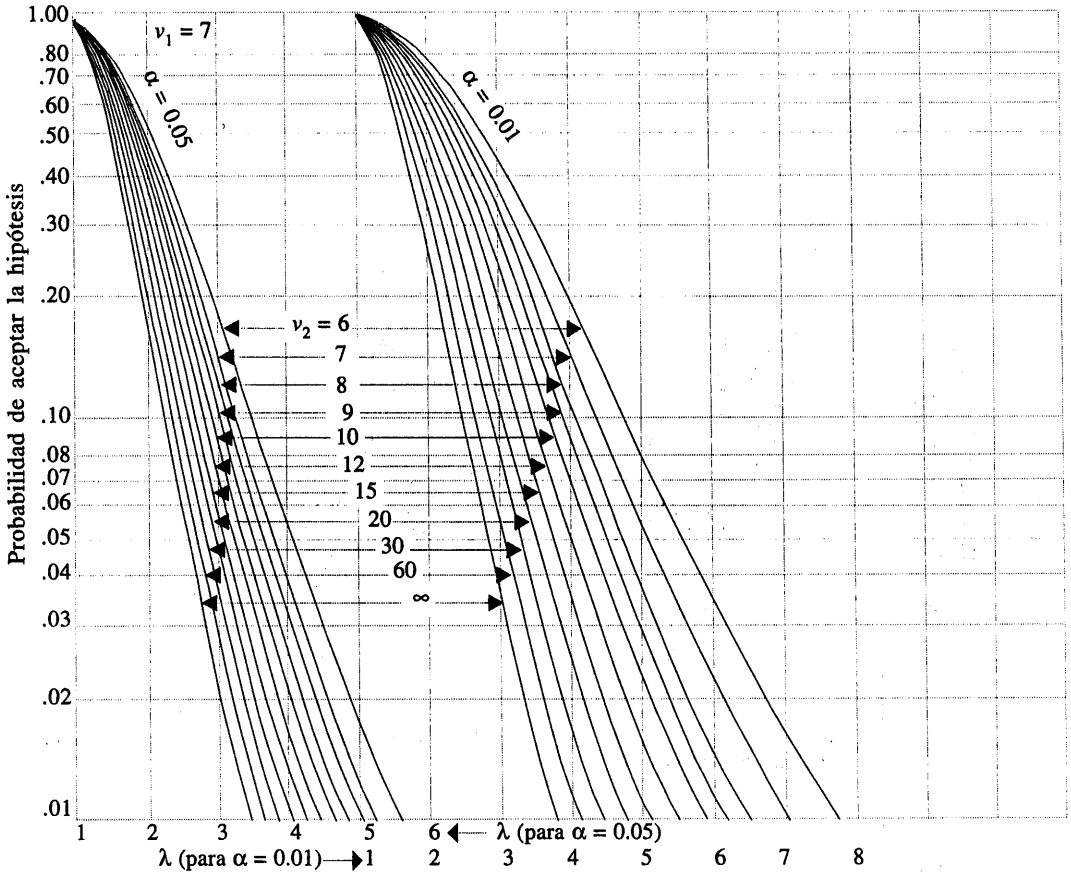


Carta VIII Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios (*continuación*)





Carta VIII Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios (*continuación*)



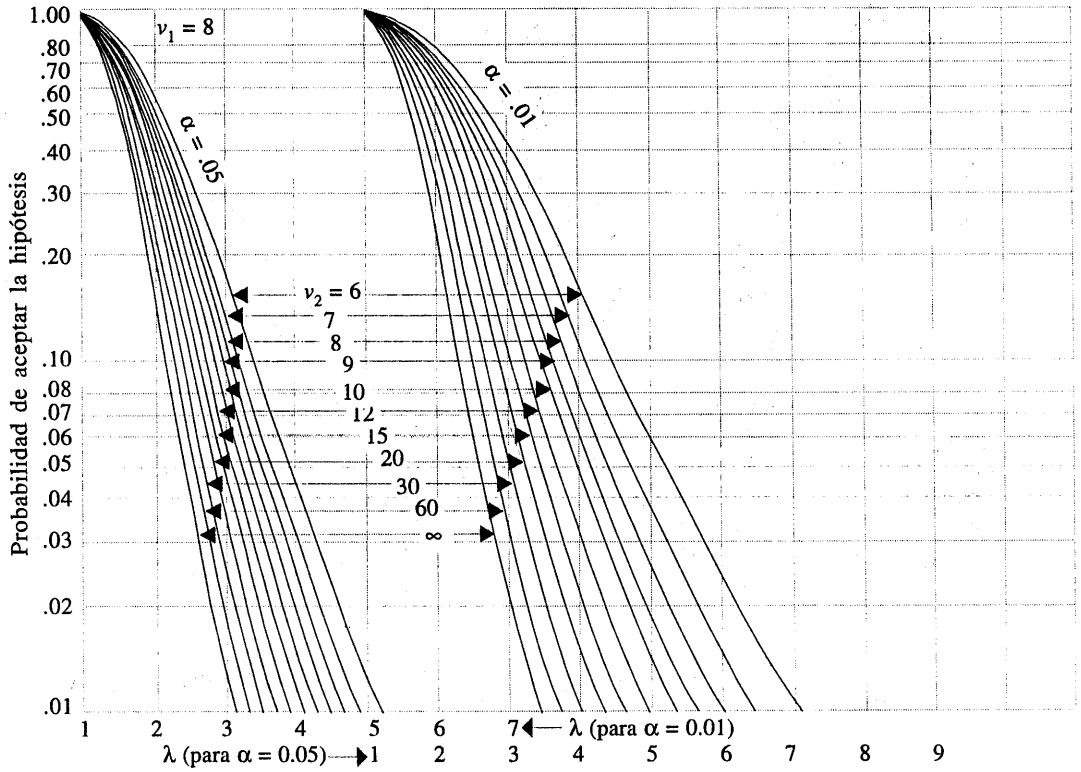


Tabla IX Valores críticos para la prueba de signos

r^*_α

n \ α	0.10 0.05	0.05 0.025	0.01 0.005	pruebas de dos colas pruebas de una cola
5	0			
6	0	0		
7	0	0		
8	1	0	0	
9	1	1	0	
10	1	1	0	
11	2	1	0	
12	2	2	1	
13	3	2	1	
14	3	2	1	
15	3	3	2	
16	4	3	2	
17	4	4	2	
18	5	4	3	
19	5	4	3	
20	5	5	3	
21	6	5	4	
22	6	5	4	
23	7	6	4	
24	7	6	5	
25	7	7	5	
26	8	7	6	
27	8	7	6	
28	9	8	6	
29	9	8	7	
30	10	9	7	
31	10	9	7	
32	10	9	8	
33	11	10	8	
34	11	10	9	
35	12	11	9	
36	12	11	9	
37	13	12	10	
38	13	12	10	
39	13	12	11	
40	14	13	11	

Tabla X Valores críticos para la prueba de rangos con signo de Wilcoxon

w^*_α

α / n^a	0.10 / 0.05	0.05 / 0.025	0.02 / 0.01	0.01 / 0.005	pruebas de dos colas / pruebas de una cola
4					
5	0				
6	2	0			
7	3	2	0		
8	5	3	1	0	
9	8	5	3	1	
10	10	8	5	3	
11	13	10	7	5	
12	17	13	9	7	
13	21	17	12	9	
14	25	21	15	12	
15	30	25	19	15	
16	35	29	23	19	
17	41	34	27	23	
18	47	40	32	27	
19	53	46	37	32	
20	60	52	43	37	
21	67	58	49	42	
22	75	65	55	48	
23	83	73	62	54	
24	91	81	69	61	
25	100	89	76	68	

[a]Si $n > 25$, W^+ (o W^-) tiene una distribución aproximadamente normal con media $n(n + 1)/4$ y varianza $n(n + 1)(2n + 1)/24$.

Tabla XI Valores críticos para la prueba de la suma de rangos de Wilcoxon

$w_{0.05}$

n_2 \ n_1[a]	4	5	6	7	8	9	10	11	12	13	14	15
4	10											
5	11	17										
6	12	18	26									
7	13	20	27	36								
8	14	21	29	38	49							
9	15	22	31	40	51	63						
10	15	23	32	42	53	65	78					
11	16	24	34	44	55	68	81	96				
12	17	26	35	46	58	71	85	99	115			
13	18	27	37	48	60	73	88	103	119	137		
14	19	28	38	50	63	76	91	106	123	141	160	
15	20	29	40	52	65	79	94	110	127	145	164	185
16	21	31	42	54	67	82	97	114	131	150	169	
17	21	32	43	56	70	84	100	117	135	154		
18	22	33	45	58	72	87	103	121	139			
19	23	34	46	60	74	90	107	124				
20	24	35	48	62	77	93	110					
21	25	37	50	64	79	95						
22	26	38	51	66	82							
23	27	39	53	68								
24	28	40	55									
25	28	42										
26	29											
27												
28												

[a]Si n_1 y $n_2 > 8$, W_1 tiene una distribución aproximadamente normal con media $\frac{1}{2}n_1(n_1 + n_2 + 1)$ y varianza $n_1n_2(n_1 + n_2 + 1)/12$.

Tabla XI Valores críticos para la prueba de la suma de rangos de Wilcoxon (*continuación*)

$w_{0.01}$

n_2 \ n_1	4	5	6	7	8	9	10	11	12	13	14	15
5		15										
6	10	16	23									
7	10	17	24	32								
8	11	17	25	34	43							
9	11	18	26	35	45	56						
10	12	19	27	37	47	58	71					
11	12	20	28	38	49	61	74	87				
12	13	21	30	40	51	63	76	90	106			
13	14	22	31	41	53	65	79	93	109	125		
14	14	22	32	43	54	67	81	96	112	129	147	
15	15	23	33	44	56	70	84	99	115	133	151	171
16	15	24	34	46	58	72	86	102	119	137	155	
17	16	25	36	47	60	74	89	105	122	140		
18	16	26	37	49	62	76	92	108	125			
19	17	27	38	50	64	78	94	111				
20	18	28	39	52	66	81	97					
21	18	29	40	53	68	83						
22	19	29	42	55	70							
23	19	30	43	57								
24	20	31	44									
25	20	32										
26	21											
27												
28												

Tabla XII Factores para construir cartas de control de variables

	Factor para los límites de control					
	Carta $\bar{X}$			Carta R		
n^a	A_1	A_2	d_2	D_3	D_4	n
2	3.760	1.880	1.128	0	3.267	2
3	2.394	1.023	1.693	0	2.575	3
4	1.880	.729	2.059	0	2.282	4
5	1.596	.577	2.326	0	2.115	5
6	1.410	.483	2.534	0	2.004	6
7	1.277	.419	2.704	.076	1.924	7
8	1.175	.373	2.847	.136	1.864	8
9	1.094	.337	2.970	.184	1.816	9
10	1.028	.308	3.078	.223	1.777	10
11	.973	.285	3.173	.256	1.744	11
12	.925	.266	3.258	.284	1.716	12
13	.884	.249	3.336	.308	1.692	13
14	.848	.235	3.407	.329	1.671	14
15	.816	.223	3.472	.348	1.652	15
16	.788	.212	3.532	.364	1.636	16
17	.762	.203	3.588	.379	1.621	17
18	.738	.194	3.640	.392	1.608	18
19	.717	.187	3.689	.404	1.596	19
20	.697	.180	3.735	.414	1.586	20
21	.679	.173	3.778	.425	1.575	21
22	.662	.167	3.819	.434	1.566	22
23	.647	.162	3.858	.443	1.557	23
24	.632	.157	3.895	.452	1.548	24
25	.619	.153	3.931	.459	1.541	25

[a] $n > 25$: $A_1 = 3/\sqrt{n}$, donde n = número de observaciones de la muestra.

Apéndice B

Material técnico complementario

I. TÉCNICAS DE CONTEO

En varios de los ejemplos del capítulo 3, es sencillo determinar el número de resultados de cada evento. En ejemplos más complicados, determinar el número de resultados que comprenden el espacio muestral (o un evento) se torna más difícil. Al asociar probabilidades con eventos, es importante conocer el número de resultados tanto de un evento como del espacio muestral. Es posible aplicar algunas reglas sencillas para simplificar los cálculos.

En el ejemplo 3-7, un fabricante de automóviles ofrece vehículos equipados con opciones seleccionadas. Cada vehículo se pide

con o sin transmisión automática

con o sin aire acondicionado

con una de tres opciones de un sistema estéreo

con uno de cuatro colores exteriores

En el diagrama de árbol de la figura 3-7 se describe el espacio muestral de todos los tipos posibles de vehículos. El tamaño del espacio muestral es igual al número de ramas en el último nivel del árbol y esta cantidad es igual a $2 \times 2 \times 3 \times 4 = 48$. Esto lleva al siguiente resultado.

Regla de multiplicación

Si una operación puede describirse como una secuencia de k pasos, y

si el número de formas en que puede completarse el paso 1 es n_1, y

si el número de formas en que puede completarse el paso 2 es n_2 para cada forma de completar el paso 1, y

si el número de formas en que puede completarse el paso 3 es n_3 para cada forma de completar el paso 2, y así sucesivamente,

entonces el número total de formas en que puede completarse la operación es

$$n_1 \times n_2 \times \cdots \times n_k$$

EJEMPLO I-1

En el diseño de la cubierta de una caja de velocidades pueden usarse cuatro tipos diferentes de sujetadores, pernos de tres longitudes diferentes, y tres localizaciones diferentes de los pernos. Por la regla de multiplicación, hay $4 \times 3 \times 3 = 36$ diseños posibles diferentes.

Otro cálculo útil es el número de arreglos (ordenaciones) de los elementos de un conjunto. Considérese un conjunto de elementos, por ejemplo $S = \{a, b, c\}$. Una **permutación** de los elementos es un arreglo u ordenación de los elementos. Así, *abc, acb, bac, bca, cab* y *cba* son todas las permutaciones de los elementos de S.

El número de **permutaciones** (arreglos) de n elementos diferentes se denota como $n!$ y

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \qquad \text{(I-1)}$$

Este resultado se sigue de la regla de multiplicación. Una permutación puede construirse seleccionando entre los n elementos el que se colocará en la primera posición del arreglo, después seleccionando de los $n - 1$ elementos restantes el que ocupará la segunda posición, después seleccionando de los $n - 2$ elementos restantes el que ocupará la tercera posición, etcétera. A las permutaciones como éstas en ocasiones se les llama permutaciones lineales.

En algunas situaciones, sólo es importante conocer el número de arreglos de algunos de los elementos de un conjunto. El resultado siguiente también se obtiene de la regla de multiplicación.

El número de permutaciones (arreglos) de un subconjunto de r elementos seleccionados de un conjunto de n elementos diferentes es

$$P_r^n = n\times(n-1)\times(n-2)\times\cdots\times(n-r+1) = \frac{n!}{(n-r)!} \qquad \text{(I-2)}$$

EJEMPLO I-2

Una tarjeta de circuitos impresos tiene ocho sitios diferentes en los que puede instalarse un componente. Si en la tarjeta deben instalarse cuatro componentes diferentes, ¿cuántos diseños diferentes son posibles?

Cada diseño consiste en seleccionar uno de los ocho sitios para el primer componente, uno de los siete sitios restantes para el segundo componente, uno de los seis sitios restantes para el tercer componente y uno de los cinco sitios restantes para el cuarto componente. Por lo tanto,

$$P_4^8 = 8\times7\times6\times5 = \frac{8!}{4!} = 1\,680 \quad \text{diseños diferentes son posible.}$$

En ocasiones el interés se encuentra en contar el número de los arreglos u ordenaciones de objetos que no son todos diferentes. El resultado siguiente es un cálculo general útil.

El número de arreglo de $n = n_1 + n_2 + \cdots + n_r$ objetos de los cuales
n_1 son de un tipo,
n_2 son de un segundo tipo,
n_3 son de un tercer tipo,
..., y
n_r son de un r-ésimo tipo,

$$\frac{n!}{n_1!\,n_2!\,n_3!\,\ldots\,n_r!} \qquad \text{(I-3)}$$

EJEMPLO I-3

Considérese una operación de maquinado en la que es necesario taladrar agujeros de dos diámetros diferentes y hacer dos muescas de tamaños diferentes en una pieza de metal laminado. La operación de taladrado se denota como d y la operación de hacer las muescas como n. Para determinar el programa de trabajo de un taller de maquinado, el responsable podría centrar su interés en el número posible de secuencias diferentes de las cuatro operaciones. El nú-

mero posible de secuencias para dos operaciones de taladrado y dos operaciones de corte de muescas es

$$\frac{4!}{2!\,2!} = 6$$

Las seis secuencias se resumen con facilidad: *ddnn, dndn, dnnd, nddn, ndnd, nndd.*

EJEMPLO I-4

Una pieza se etiqueta imprimiendo cuatro líneas gruesas, tres líneas medianas y dos líneas delgadas. Si cada ordenación de las nueve líneas representa una etiqueta diferente, ¿cuántas etiquetas diferentes pueden generarse utilizando este esquema?

Por la ecuación I-3, el número de etiquetas posibles para las piezas es

$$\frac{9!}{4!\,3!\,2!} = 2\ 520$$

Otro problema de conteo que podría ser de interés es el número de subconjuntos de r elementos que pueden seleccionarse de un conjunto de n elementos. En este caso, el orden no es importante. Cada subconjunto de r elementos puede indicarse enlistando los elementos del conjunto y marcando cada elemento que va a incluirse en el subconjunto con un asterisco "*". Por lo tanto, cada arreglo de r asteriscos "*" y $n - r$ elementos no marcados indica un subconjunto diferente. Por la ecuación I-3, se obtiene el siguiente resultado.

El número de subconjuntos de tamaño r que pueden seleccionarse de un conjunto de n elementos se denota como $\binom{n}{r}$ o C_r^n y

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \qquad \text{(I-4)}$$

EJEMPLO I-5

Una tarjeta de circuitos impresos tiene ocho sitios diferentes en los que puede instalarse un componente. Si van a colocarse cinco componentes idénticos en la tarjeta, ¿cuántos diseños diferentes son posibles?

Cada diseño es un subconjunto de los ocho sitios que van a contener los componentes. Por la ecuación I-4, el número posible de diseños es

$$\frac{8!}{5!\,3!} = 56$$

En el ejemplo siguiente se usa la regla de multiplicación en combinación con la ecuación I-4 para encontrar la respuesta de una pregunta más complicada, pero común.

EJEMPLO I-6

Un depósito de 50 piezas manufacturadas contiene tres piezas defectuosas y 47 piezas no defectuosas. Se selecciona una muestra de seis piezas del lote de 50. Las piezas seleccionadas no son reemplazadas. Es decir, cada parte puede seleccionarse una sola vez y la muestra es un subconjunto de las 50 piezas. ¿Cuántas muestras diferentes de tamaño seis hay que contienen exactamente dos piezas defectuosas?

Un subconjunto que contiene exactamente dos piezas defectuosas puede formarse escogiendo primero dos de las tres piezas defectuosas. Utilizando la ecuación I-4, este paso puede completarse en

$$\frac{3!}{2!\,1!} = 3 \text{ formas diferentes}$$

Después, el segundo paso es seleccionar las cuatro partes restantes entre las 47 piezas satisfactorias en el depósito. El segundo paso puede completarse en

$$\frac{47!}{4!\,43!} = 178\,365 \text{ formas diferentes}$$

Por lo tanto, por la regla de multiplicación, el número de subconjuntos de tamaño seis que contienen exactamente dos artículos defectuosos es

$$3 \times 178\,365 = 535\,095$$

Como un cálculo adicional, se encuentra que el número total de subconjuntos diferentes de tamaño seis es

$$\frac{50!}{6!\,44!} = 15\,890\,700$$

Las piezas del ejemplo I-6 se eligen sin reemplazar previamente las piezas seleccionadas. En consecuencia, cada pieza puede seleccionarse una sola vez y la muestra es un subconjunto del conjunto de todas las piezas. A este método de selección se le llama con frecuencia muestreo **sin reemplazo**. En el capítulo 3 se estableció que la selección de una pieza es aleatoria si cada pieza es igualmente posible. Cuando se selecciona más de una pieza, se dice que el muestreo es **aleatorio** si cada subconjunto de piezas es igualmente posible de ser seleccionado.

EJEMPLO I-7

En el ejemplo I-6, si el muestreo es aleatorio, entonces cada subconjunto de seis piezas es igualmente posible. Por lo tanto, la probabilidad de que una muestra contenga exactamente dos piezas defectuosas es

$$\frac{535\,095}{15\,890\,700} = 0.034$$

Obsérvese que el ejemplo I-7 ilustra la distribución hipergeométrica.

Ejercicios del Apéndice I

I-1. En el pedido de un asistente digital personal puede especificarse cualquiera de cinco tamaños de memoria, cualquiera de tres tipos de pantalla, cualquiera de cuatro tamaños del disco duro, y puede incluir o no un lápiz óptico. ¿Cuántos pedidos diferentes del sistema pueden hacerse?

I-2. En una operación de manufactura, una pieza se produce por maquinado, pulido y pintado. Si hay tres herramientas de maquinado, cuatro herramientas de pulido y tres herramientas de pintado, ¿cuántas rutas diferentes (las cuales constan de maquinado, seguido por el pulido y después por el pintado) para una pieza son posibles?

I-3. En los nuevos diseños de un tanque de tratamiento de aguas negras se han propuesto tres formas posibles, cuatro tamaños posibles, tres sitios para las válvulas de entrada y cuatro sitios para las válvulas de salida. ¿Cuántos diseños diferentes del producto son posibles?

I-4. Un proceso de manufactura consta de 10 operaciones diferentes que pueden completarse en cualquier orden. ¿Cuántas secuencias diferentes de producción son posibles?

I-5. Un proceso de manufactura consta de 10 operaciones. Sin embargo, es necesario completar cinco operaciones de maquinado antes de que pueda iniciarse cualquiera de las operaciones de ensamblaje. Dentro de cada conjunto de cinco, las operaciones pueden completarse en cualquier orden. ¿Cuántas secuencias diferentes de producción son posibles?

I-6. En una operación de metal laminado se requieren tres muescas y cuatro dobleces. Si las operaciones pueden realizarse en cualquier orden, ¿cuántas formas diferentes de completar la manufactura son posibles?

I-7. Un lote de 140 chips semiconductores se inspecciona escogiendo una muestra de cinco chips. Suponga que 10 de los chips no cumplen con los requerimientos del cliente.

a) ¿Cuántas muestras diferentes son posibles?

b) ¿Cuántas muestras de cinco chips contienen exactamente uno no satisfactorio?

c) ¿Cuántas muestras de cinco chips contienen al menos uno no satisfactorio?

I-8. En la disposición de una tarjeta de circuitos impresos para un producto electrónico hay 12 sitios diferentes donde pueden colocarse chips.

a) Si van a colocarse cinco tipos diferentes de chips en la tarjeta, ¿cuántas disposiciones diferentes son posibles?

b) Si los cinco chips que se colocan en la tarjeta son del mismo tipo, ¿cuántas disposiciones diferentes son posibles?

I-9. En el laboratorio de análisis de muestras de un proceso químico se analizan diariamente cinco muestras del proceso. Además, se analiza una prueba de control dos veces cada día para verificar la calibración de los instrumentos del laboratorio.

a) ¿Cuántas secuencias diferentes de las muestras del proceso y de control son posibles cada día? Suponga que las cinco muestras del proceso se consideran idénticas y que las dos muestras de control se consideran idénticas.

b) ¿Cuántas secuencias diferentes de las muestras del proceso y de control son posibles si se considera que las cinco muestras del proceso son diferentes y que las dos muestras de control son idénticas?

c) Para la misma situación que en el inciso b), ¿cuántas secuencias son posibles si la primera prueba de cada día debe ser una muestra de control?

I-10. En el diseño de un producto electromecánico, siete componentes diferentes deben apilarse en un contenedor cilíndrico con capacidad para 12 componentes de tal manera que se minimiza el impacto de los golpes. Uno de los extremos del contenedor está diseñado como fondo y el otro como la parte superior.

a) ¿Cuántos diseños diferentes son posibles?

b) Si los siete componentes son idénticos, ¿cuántos diseños diferentes son posibles?

c) Si los siete componentes constan de tres de un tipo y cuatro de otro tipo, ¿cuántos diseños diferentes son posibles? (Más difícil.)

I-11. En el diseño de un sistema de comunicaciones se consideraron las siguientes cuestiones.

a) ¿Cuántos prefijos telefónicos de tres dígitos que se usan para representar un área geográfica particular (por ejemplo un código de área) pueden crearse con los dígitos del 0 al 9?

b) Como en el inciso a), ¿cuántos prefijos telefónicos de tres dígitos que no empiecen con 0 o 1, pero que contengan 0 o 1 como el dígito de en medio son posibles?

c) ¿Cuántos prefijos telefónicos de tres dígitos en los que ninguno de los dígitos aparezca más de una vez en cada prefijo son posibles?

I-12. Un byte es una secuencia de ocho bits y cada bit es 0 o 1.

a) ¿Cuántos bytes diferentes son posibles?

b) Si el primer bit de un byte es un verificador de paridad, es decir, el primer bit está determinado por los siete bits restantes, ¿cuántos bytes diferentes son posibles?

I-13. En una planta química se usan 24 tanques para almacenar el producto final. Se seleccionan cuatro tanques al azar y sin reemplazo. Suponga que seis de los tanques contienen material en el que la viscosidad excede los requerimientos del cliente.

a) ¿Cuál es la probabilidad de que exactamente un tanque de la muestra contenga material con viscosidad elevada?

b) ¿Cuál es la probabilidad de que al menos uno de los tanques de la muestra contenga material con viscosidad elevada?

c) Además de los seis tanques con niveles de viscosidad elevada, cuatro tanques diferentes contienen material con impurezas altas. ¿Cuál es la probabilidad de que exactamente uno de los tanques de la muestra contenga material con viscosidad elevada y de que exactamente uno de los tanques en la muestra, contenga material con impurezas altas?

I-14. Se verifica que las partes de plástico que se producen en una operación de moldeo por inyección se ajustan a las especificaciones. Cada herramienta contiene 12 cavidades en las que se producen las partes, las cuales caen en una transportadora cuando se abre la prensa. Un inspector escoge al azar tres partes de entre las 12. Dos cavidades están afectadas por una falla de temperatura que resulta en partes que no cumplen con las especificaciones.

a) ¿Cuál es la probabilidad de que el inspector encuentre exactamente una pieza no satisfactoria?

b) ¿Cuál es la probabilidad de que el inspector encuentre al menos una parte no satisfactoria?

I-15. Un depósito de 50 piezas contiene cinco que están defectuosas. Se selecciona al azar una muestra de dos, sin reemplazo.

a) Determine la probabilidad de que las dos piezas de la muestra estén defectuosas calculando una probabilidad condicional como en el capítulo 2.

b) Determine la respuesta del inciso a) aplicando el enfoque del subconjunto que se describió en esta sección.

II. FUNCIÓN GENERADORA DE MOMENTOS

Suponga que X es una variable aleatoria con media μ. A lo largo de este libro se ha utilizado la idea del valor esperado de la variable aleatoria X y, de hecho, $E(X) = \mu$. Suponga ahora que el interés se encuentra en el valor esperado de una función particular de X, por ejemplo, $g(X) = X^r$. Al valor esperado de esta función, o $E[g(X)] = E(X^r)$, se le llama el momento r-ésimo alrededor del origen de la variable aleatoria X, el cual se denotará por μ'_r.

Definición

El **momento *r*-ésimo alrededor del origen** de la variable aleatoria X es

$$\mu'_r = E(X^r) = \begin{cases} \sum_x x^r f(x), & X \text{ discreta} \\ \int_{-\infty}^{\infty} x^r f(x)\,dx, & X \text{ continua} \end{cases} \qquad \text{(II-1)}$$

Obsérvese que el primer momento alrededor del origen es la media, es decir, $E(X) = \mu'_1 = \mu$. Además, puesto que el segundo momento alrededor del origen es $E(X^2) = \mu'_2$, la varianza de una variable aleatoria pueden expresarse en términos de momentos al origen como sigue:

$$\begin{aligned} \sigma^2 &= E(X^2) - [E(X)]^2 \\ &= \mu'_2 - \mu^2 \end{aligned}$$

Muchas veces el momento de una variable aleatoria puede determinarse directamente a partir de la definición de la ecuación II-1, pero existe un procedimiento alternativo que hace uso de una función especial que con frecuencia es útil.

Definición

La **función generadora de momentos** de la variable aleatoria X es el valor esperado de e^{tX} y se denota por $M_X(t)$. Es decir,

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x), & X \text{ discreta} \\ \int_{-\infty}^{\infty} e^{tx} f(x)\,dx, & X \text{ continua} \end{cases}$$

La función generadora de momentos $M_X(t)$ existirá solamente si la operación suma o la integral de la definición anterior converge. Si la función generadora de momentos de una variable aleatoria existe, puede usarse para obtener todos los momentos al origen de la variable aleatoria.

Teorema II-1

Sea X una variable aleatoria con función generadora de momentos $M_X(t)$. Entonces

$$\mu'_r = \left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0}$$

Demostración Suponiendo que es posible tomar derivadas dentro de los signos de la operación suma e integral, entonces

$$\frac{d^r M_X(t)}{dt^r} = \begin{cases} \sum_x x^r e^{tx} f(x), & X \text{ discreta} \\ \int_{-\infty}^{\infty} x^r e^{tx} f(x)\,dx, & X \text{ continua} \end{cases} \tag{II-2}$$

Ahora bien, si se hace $t = 0$ en la ecuación II-2, se encuentra que

$$\left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0} = E(X^r)$$

EJEMPLO II-1

Suponga que X tiene una distribución binominal, es decir

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$$
$$= 0, \qquad \text{en caso contrario}$$

Encuentre la función generadora de momentos y úsela para verificar que la media y la varianza de la variable aleatoria binomial son $\mu = np$ y $\sigma^2 = np(1-p)$.

Por la definición de función generadora de momentos, se tiene

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$
$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}$$

Esta última operación suma es la expansión binomial de $[pe^t + (1-p)]^n$, de donde

$$M_X(t) = [pe^t + (1-p)]^n$$

Al sacar la primera y la segunda derivada, se obtiene

$$M'_X(t) = \frac{dM_X(t)}{dt} = npe^t[1 + p(e^t - 1)]^{n-1}$$

y

$$M''_X(t) = \frac{d^2M_X(t)}{dt^2} = npe^t(1 - p + npe^t)[1 + p(e^t - 1)]^{n-2}$$

Si se hace $t = 0$ en $M'_X(t)$, se obtiene

$$M'_X(t)|_{t=0} = \mu'_1 = \mu = np$$

que es la media de la variable aleatoria binominal X. Si ahora se hace $t = 0$ en $M'_X(t)$, entonces

$$M'_X(t)|_{t=0} = \mu'_2 = np(1 - p + np)$$

Por lo tanto, la varianza de la variable aleatoria binomial es

$$\begin{aligned}\sigma^2 &= \mu'_2 - \mu^2 \\ &= np(1 - p + np) - (np)^2 \\ &= np - np^2 \\ &= np(1 - p)\end{aligned}$$

EJEMPLO II-2

Encuentre la función generadora de momentos de la variable aleatoria normal y úsela para demostrar que la media y la varianza de esta variable aleatoria son μ y σ^2, respectivamente.

La función generadora de momentos es

$$\begin{aligned}M_X(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-[x^2 - 2(\mu + t\sigma^2)x + \mu^2]/(2\sigma^2)} dx\end{aligned}$$

Si se completa el trinomio cuadrado perfecto en el exponente, se tiene

$$x^2 - 2(\mu + t\sigma^2)x + \mu^2 = [x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4$$

y entonces

$$M_X(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\{[x-(\mu+t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4\}/(2\sigma^2)} dx$$

$$= e^{\mu t + \sigma^2 t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[x-(\mu+t\sigma^2)]^2/\sigma^2} dx$$

Sea $u = [x - (\mu + t\sigma^2)]/\sigma$. Entonces $dx = \sigma\, du$, y la última expresión anterior pasa a ser

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Ahora bien, la integral es el área total bajo una curva de densidad normal estándar, que es 1, por lo que la función generadora de momentos de una variable aleatoria normal es

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$$

Al derivar dos veces esta función con respecto a t y haciendo $t = 0$ en el resultado, se encuentra

$$\left.\frac{dM_X(t)}{dt}\right|_{t=0} = \mu_1' = \mu$$

$$\left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \mu_2' = \sigma^2 = \mu^2$$

Por lo tanto, la varianza de la variable aleatoria normal es

$$\sigma^2 = \mu_2' - \mu^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

Las funciones generadoras de momentos poseen importantes y útiles propiedades. Una de las más importantes es la **propiedad de unicidad**. Es decir, la función generadora de momentos de una variable aleatoria es única cuando existe, de tal modo que si se tienen dos variables aleatorias X y Y, por ejemplo, con funciones generadoras de momentos $M_X(t)$ y $M_Y(t)$, entonces si $M_X(t) = M_Y(t)$ para todos los valores de t, tanto X como Y tienen la misma distribución de probabilidad. En la página siguiente se resumen otras propiedades útiles de la función generadora de momentos.

Teorema II-2

Si X es una variable aleatoria y a es una constante, entonces

$$M_{X+a}(t) = e^{at} M_X(t)$$

Demostración

$$\begin{aligned} M_{X+a}(t) &= E\left[e^{t(X+a)}\right] \\ &= e^{at} E\left(e^{tX}\right) \\ &= e^{at} M_X(t) \end{aligned}$$

Teorema II-3

Si X es una variable aleatoria y a es una constante, entonces

$$M_{aX}(t) = M_X(at)$$

Demostración

$$\begin{aligned} M_{aX}(t) &= E\left[e^{t(aX)}\right] \\ &= E\left(e^{(at)X}\right) \\ &= M_X(at) \end{aligned}$$

Teorema II-4

Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes con funciones generadoras de momentos $M_{X_1}(t), M_{X_2}(t), \ldots, M_{X_n}(t)$, respectivamente, y si $Y = X_1 + X_2 + \cdots + X_n$, entonces la función generadora de momentos de Y es

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \cdots \cdot M_{X_n}(t)$$

Demostración Para el caso en que las X son variables aleatorias continuas

$$M_Y(t) = E(e^{tY}) = E\left[e^{t(X_1+X_2+\cdots+X_n)}\right]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t(x_1+x_2+\ldots+x_n)} f(x_1, x_2, \ldots, x_n)\, dx_1, dx_2, \ldots, dx_n$$

Puesto que las X son independientes,

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \ldots \cdot f_{X_n}(x_n)$$

y puede escribirse

$$M_Y(t) = \int_{-\infty}^{\infty} e^{tx_1} f_{X_1}(x_1)\, dx_1 \int_{-\infty}^{\infty} e^{tx_2} f_{X_2}(x_2)\, dx_2 \cdots \int_{-\infty}^{\infty} e^{tx_n} f_{X_n}(x_n)\, dx_n$$
$$= M_{X_1}(t) \cdot M_{X_2}(t) \cdot \ldots \cdot M_{X_n}(t)$$

Para el caso en que las X son discretas, se usaría el mismo enfoque reemplazando las integrales con operaciones suma.

El teorema II-4 es de particular utilidad. En muchas situaciones es necesario encontrar la distribución de la suma de dos o más variables aleatorias independientes, y con frecuencia el problema se simplifica bastante con este teorema. Esto se ilustra en el ejemplo siguiente.

EJEMPLO II-3

Suponga que X_1 y X_2 son dos variables aleatorias de Poisson independientes con parámetros λ_1 y λ_2, respectivamente. Encuentre la distribución de probabilidad de $Y = X_1 + X_2$.

La función generadora de momentos de una variable aleatoria de Poisson con parámetro λ es

$$M_X(t) = e^{\lambda(e^t-1)}$$

por lo que las funciones generadoras de momentos de X_1 y X_2 son $M_{X_1}(t) = e^{\lambda_1(e^t-1)}$ y $M_{X_2}(t) = e^{\lambda_2(e^t-1)}$, respectivamente. Utilizando los resultados del teorema II-4, se encuentra que la función generadora de momentos de $Y = X_1 + X_2$ es

$$M_Y(t) = M_{X_1}(t) M_{X_2}(t)$$
$$= e^{\lambda_1(e^t-1)} e^{\lambda_2(e^t-1)}$$
$$= e^{(\lambda_1+\lambda_2)(e^t-1)}$$

expresión que se identifica como la función generadora de momentos de una variable aleatoria de Poisson con parámetro $\lambda_1 + \lambda_2$. Por lo tanto, se ha demostrado que la suma de dos variables aleatorias de Poisson independientes con parámetros λ_1 y λ_2 es una variable aleatoria de Poisson con parámetro igual a la suma de los dos parámetros $\lambda_1 + \lambda_2$.

EJERCICIOS

II-1. Una variable aleatoria X tiene la distribución uniforme discreta

$$f(x)=\begin{cases}\dfrac{1}{m}, & x=1,2,\ldots,m\\ 0, & \text{en caso contrario}\end{cases}$$

a) Demuestre que la función generadora de momentos es

$$M_X(t)=\frac{e^t(1-e^{tm})}{m(1-e^t)}$$

b) Use $M_X(t)$ para encontrar la media y la varianza de X.

II-2. Una variable aleatoria X tiene la distribución de Poisson

$$f(x)=\begin{cases}\dfrac{e^{-\lambda}\lambda^x}{x!}, & x=0,1,\ldots\\ 0, & \text{en caso contrario}\end{cases}$$

a) Demuestre que la función generadora de momentos es

$$M_X(t)=e^{\lambda(e^t-1)}$$

b) Use $M_X(t)$ para encontrar la media y la varianza de la variable aleatoria de Poisson.

II-3. La variable aleatoria geométrica X tiene la distribución de probabilidad

$$f(x)=\begin{cases}(1-p)^{x-1}p, & x=1,2,\ldots\\ 0, & \text{en caso contrario}\end{cases}$$

a) Demuestre que la función generadora de momentos es

$$M_X(t)=\frac{pe^t}{1-(1-p)e^t}$$

b) Use $M_X(t)$ para encontrar la media y la varianza de X.

II-4. La variable aleatoria ji-cuadrada con k grados de libertad tiene la función generadora de momentos $M_X(t) = (1 - 2t)^{-k/2}$. Suponga que X_1 y X_2 son variables aleatorias ji-cuadrada independientes con k_1 y k_2 grados de libertad, respectivamente. ¿Cuál es la distribución de $Y = X_1 + X_2$?

II-5. Una variable aleatoria continua X tiene la siguiente distribución de probabilidad:

$$f(x)=\begin{cases}4xe^{-2x}, & x>0\\ 0, & \text{en caso contrario}\end{cases}$$

a) Encuentre la función generadora de momentos de X.

b) Encuentre la media y la varianza de X.

II-6. La variable aleatoria uniforme y continua X tiene la función de densidad

$$f(x)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \alpha\le x\le\beta\\ 0, & \text{en caso contrario}\end{cases}$$

a) Demuestre que la función generadora de momentos es

$$M_X(t)=\frac{e^{t\beta}-e^{t\alpha}}{t(\beta-\alpha)}$$

b) Use $M_X(t)$ para encontrar la media y la varianza de X.

II-7. Una variable aleatoria X tiene la distribución exponencial

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & \text{en caso contrario}\end{cases}$$

a) Demuestre que la función generadora de momentos de X es

$$M_X(t)=\left(1-\frac{t}{\lambda}\right)^{-1}$$

b) Encuentre la media y la varianza de X.

II-8. Una variable aleatoria X tiene la distribución gamma

$$f(x)=\begin{cases}\dfrac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}, & x>0\\ 0, & \text{en caso contrario}\end{cases}$$

a) Demuestre que la función generadora de momentos de X es

$$M_X(t)=\left(1-\frac{t}{\lambda}\right)^{-r}$$

b) Encuentre la media y la varianza de X.

II-9. Sean $X_1, X_2, \ldots, X_r$ variables aleatorias exponenciales independientes con parámetro λ.

a) Encuentre la función generadora de momentos de $Y = X_1 + X_2 + \cdots + X_r$.

b) ¿Cuál es la distribución de la variable aleatoria Y?

Sugerencia: Use los resultados del ejercicio II-8.

II-10. Suponga que X_i tiene una distribución normal con media μ_i y varianza σ_i^2, $i = 1, 2$. Sean X_1 y X_2 independientes.

a) Encuentre la función generadora de momentos de $Y = X_1 + X_2$.

b) ¿Cuál es la distribución de la variable aleatoria Y?

II-11. Demuestre que la función generadora de momentos de la variable aleatoria ji-cuadrada con k grados de libertad es $M_X(t) = (1 - 2t)^{-k/2}$. Demuestre que la media y la varianza de esta variable aleatoria son k y $2k$, respectivamente.

II-12. a) Demuestre que al hacer la expansión de e^{tX} en una serie de potencias y al tomar los esperados término a término, la función generadora de momentos puede escribirse como

$$M_X(t) = E(e^{tX})$$
$$= 1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \cdots$$
$$+ \mu_r' \frac{t^r}{r!} + \cdots$$

Por tanto, el coeficiente de $t^r/r!$ en esta expansión es μ_1', el r-ésimo momento al origen.

b) Considérese la variable aleatoria gamma del ejercicio II-8 anterior. Escriba la expansión de una serie de potencias de $M_X(t)$.

c) Encuentre μ_1' y μ_2' utilizando los resultados de los incisos a) y b). ¿Con este enfoque se obtienen los mismos resultados encontrados para la media y la varianza de la variable aleatoria gamma del ejercicio II-8?

III. FUNCIONES DE VARIABLES ALEATORIAS

En muchas situaciones de la estadística es necesario deducir la distribución de probabilidad de una función de una o más variables aleatorias. En esta sección, se presentan algunos resultados que son de utilidad para resolver este problema.

Suponga que X es una variable aleatoria discreta con distribución de probabilidad $f_X(x)$. Sea $Y = h(X)$ una función de X que define una transformación uno a uno entre los valores de X y Y, y quiere encontrarse la distribución de probabilidad de Y. Por transformación uno a uno se entiende que cada valor de x se relaciona con uno y sólo un valor de $y = h(x)$, y que cada valor de y se relaciona con uno y sólo un valor de x, por ejemplo, $x = u(y)$, donde $u(y)$ se encuentra resolviendo $y = h(x)$ para determinar x en términos de y.

Ahora bien, la variable aleatoria Y toma el valor y cuando X toma el valor $u(y)$. Por lo tanto, la distribución de probabilidad de Y es

$$f_Y(y) = P(Y = y) = P[X = u(y)] = f_X[u(y)]$$

Este resultado puede enunciarse como un teorema.

Teorema III-1

Suponga que X es una variable aleatoria **discreta** con distribución de probabilidad $f_X(x)$. Sea que $Y = h(X)$ defina una transformación uno a uno entre los valores de X y Y de tal modo que la ecuación $y = h(x)$ tiene una solución única para x en términos de y. Sea esta solución $x = u(y)$. Entonces la distribución de probabilidad de la variable aleatoria Y es

$$f_Y(y) = f_X[u(y)]$$

EJEMPLO III-1

Sea X una variable aleatoria geométrica con distribución de probabilidad

$$f_X(x) = \begin{cases} p(1-p)^{x-1}, & x = 1, 2, \ldots \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de probabilidad de $Y = X^2$.

Puesto que $X \geq 0$, la transformación es uno a uno; es decir, $y = x^2$ y $x = \sqrt{y}$. Por lo tanto, el teorema indica que la distribución de la variable aleatoria Y es

$$f_Y(y) = \begin{cases} f(\sqrt{y}) = p(1-p)^{\sqrt{y}-1}, & y = 1, 4, 9, 16, \ldots \\ 0, & \text{en caso contrario} \end{cases}$$

Suponga ahora que se tienen dos variables aleatorias discretas X_1 y X_2 con distribución de probabilidad conjunta $f_{X_1X_2}(x_1, x_2)$ y que quiere encontrarse la distribución de probabilidad conjunta $f_{Y_1Y_2}(y_1, y_2)$ de las dos nuevas variables aleatorias $Y_1 = h_1(X_1, X_2)$ y $Y_2 = h_2(X_1, X_2)$. Se supone que las funciones h_1 y h_2 definen una transformación uno a uno entre (x_1, x_2) y (y_1, y_2). Al resolver simultáneamente las ecuaciones $y_1 = h_1(x_1, x_2)$ y $y_2 = h_2(x_1, x_2)$, se obtiene la solución única $x_1 = u_1(y_1, y_2)$ y $x_2 = u_2(y_1, y_2)$. Por lo tanto, las variables aleatorias Y_1 y Y_2 toman los valores y_1 y y_2 cuando X_1 toma el valor $u_1(y_1, y_2)$ y X_2 toma el valor $u_2(y_1, y_2)$. La distribución de probabilidad conjunta de Y_1 y Y_2 es

$$\begin{aligned} f_{Y_1Y_2}(y_1, y_2) &= P(Y_1 = y_1, Y_2 = y_2) \\ &= P[X_1 = u_1(y_1, y_2), X_2 = u_2(y_1, y_2)] \\ &= f_{X_1X_2}[u_1(y_1, y_2), u_2(y_1, y_2)] \end{aligned}$$

Este resultado también se enunciará como un teorema.

Teorema III-2

Suponga que X_1 y X_2 son variables aleatorias **discretas** con distribución de probabilidad conjunta $f_{X_1X_2}(x_1, x_2)$ y sea que $Y_1 = h_1(X_1, X_2)$ y $Y_2 = h_2(X_1, X_2)$ definan una transformación uno a uno entre los puntos (x_1, x_2) y (y_1, y_2), de tal modo que las ecuaciones $y_1 = h_1(x_1, x_2)$ y $y_2 = h_2(x_1, x_2)$ tienen una solución única para x_1 y x_2 en términos de y_1 y y_2. Sea esta solución $x_1 = u_1(y_1, y_2)$ y $x_2 = u_2(y_1, y_2)$. Entonces la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$f_{Y_1Y_2}(y_1, y_2) = f_{X_1X_2}[u_1(y_1, y_2), u_2(y_1, y_2)]$$

Una aplicación muy importante del teorema III-2 es para encontrar la distribución de una variable aleatoria Y_1 que es una función de otras dos variables aleatorias X_1 y X_2. Es decir, sea $Y_1 = h_1(X_1, X_2)$, donde X_1 y X_2 son variables aleatorias discretas con distribución conjunta $f_{X_1X_2}(x_1, x_2)$. Quiere encontrarse la distribución de probabilidad de Y_1, por ejemplo, $f_{Y_1}(y_1)$. Para ello, se define una segunda función $Y_2 = h_2(X_1, X_2)$, de tal modo que se mantenga la correspondencia uno a uno entre los puntos (x_1, x_2) y (y_1, y_2), y se usa el resultado del teorema III-2 para encontrar la distribución de probabilidad conjunta de Y_1 y Y_2. Entonces la distribución de Y_1 sola se encuentra con la suma sobre los valores y_2 de esta distribución conjunta. Es decir, $f_{Y_1}(y_1)$ es la **distribución de probabilidad marginal** de Y_1, o

$$f_{Y_1}(y_1) = \sum_{y_2} f_{Y_1Y_2}(y_1, y_2)$$

EJEMPLO III-2

Considérese el caso en que X_1 y X_2 son variables aleatorias de Poisson independientes con parámetros λ_1 y λ_2, respectivamente. Se encontrará la distribución de la variable aleatoria $Y_1 = X_1 + X_2$.

La distribución conjunta de X_1 y X_2 es

$$\begin{aligned} f_{X_1X_2}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2}(x_2) \\ &= \frac{e^{-\lambda_1}\lambda_1^{x_1}}{x_1!}\,\frac{e^{-\lambda_2}\lambda_2^{x_2}}{x_2!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}\lambda_1^{x_1}\lambda_2^{x_2}}{x_1!\,x_2!}, \qquad x_1 = 0, 1, \ldots \qquad x_2 = 0, 1, \ldots \end{aligned}$$

ya que X_1 y X_2 son independientes. Para usar ahora el teorema III-2 es necesario definir una segunda función $Y_2 = h_2(X_1, X_2)$. Sea esta función $Y_2 = X_2$. Entonces las soluciones para x_1 y x_2 son $x_1 = y_1 - y_2$ y $x_2 = y_2$. Por tanto, por el teorema III-2, la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$f_{Y_1Y_2}(y_1, y_2) = \frac{e^{-(\lambda_1+\lambda_2)}\lambda_1^{(y_1-y_2)}\lambda_2^{y_2}}{(y_1-y_2)!\,y_2!}, \qquad y_1 = 0, 1, 2, \ldots \qquad y_2 = 0, 1, \ldots, y_1$$

Puesto que $x_1 \geq 0$, la transformación $x_1 = y_1 - y_2$ requiere que $x_2 = y_2$ debe ser siempre menor o igual que y_1. Por tanto, los valores de y_2 son 0, 1, ..., y_1, y la distribución de probabilidad marginal de Y_1 se obtiene como sigue:

$$f_{Y_1}(y_1) = \sum_{y_2=0}^{y_1} f_{Y_1Y_2}(y_1, y_2) = \sum_{y_2=0}^{y_1} \frac{e^{-(\lambda_1+\lambda_2)}\lambda_1^{(y_1-y_2)}\lambda_2^{y_2}}{(y_1-y_2)!\,y_2!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{y_1!} \sum_{y_2=0}^{y_1} \frac{y_1!\lambda^{(\lambda_1-\lambda_2)}\lambda_2^{y_2}}{(y_1-y_2)!\,y_2!} = \frac{e^{-(\lambda_1+\lambda_2)}}{y_1!} \sum_{y_2=0}^{y_1} \binom{y_1}{y_2} \lambda_1^{(\lambda_1-\lambda_2)}\lambda_2^{y_2}$$

La suma de esta última expresión es la expansión binomial de $(\lambda_1 + \lambda_2)^{y_1}$, de donde

$$f_{Y_1}(y_1) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^{y_1}}{y_1!}, \qquad y_1 = 0, 1, \ldots$$

Esta expresión se identifica como una distribución de Poisson con parámetro $\lambda_1 + \lambda_2$. Por lo tanto, se ha demostrado que la suma de dos variables aleatorias de Poisson independientes con parámetros λ_1 y λ_2 tiene una distribución de Poisson con parámetro $\lambda_1 + \lambda_2$.

Se considera ahora la situación en que las variables aleatorias son continuas. El teorema III-3 trata el caso en que $Y = h(X)$, con X continua, y la transformación es uno a uno.

Teorema III-3

Suponga que X es una variable aleatoria **continua** con distribución de probabilidad $f_X(x)$. La función $Y = h(X)$ es una transformación uno a uno entre los valores de Y y X de tal modo que la ecuación $y = h(x)$ tiene una solución única para x en términos de y. Sea esta solución $x = u(y)$. La distribución de probabilidad de Y es

$$f_Y(y) = f_X[u(y)]|J|$$

donde a $J = u'(y)$ se le llama el **jacobiano** de la transformación y se usa el valor absoluto de J.

Demostración Sea $y = h(x)$ una función creciente de x. Ya que

$$P(Y \le a) = P[X \le u(a)]$$
$$= \int_{-\infty}^{u(a)} f_X(x)dx$$

Si se cambia la variable de integración de x a y utilizando $x = u(y)$, se obtiene $dx = u'(y)\,dy$ y entonces

$$P(Y \le a) = \int_{-\infty}^{a} f_X[u(y)]u'(y)dx$$

Puesto que la integral da la probabilidad de que $Y \le a$ para todos los valores de a contenidos en el conjunto de los valores apropiados de y, entonces $f_X[u(y)]\,u'(y)$ debe ser la distribución de probabilidad de Y. Por lo tanto, la distribución de probabilidad de Y es

$$f_Y(y) = f_X[u(y)]u'(y) = f_X[u(y)]J$$

Si la función $y = h(x)$ es una función decreciente de x, es válido un razonamiento similar.

EJEMPLO III-3

Sea X una variable aleatoria continua con distribución de probabilidad

$$f_X(x) = \begin{cases} \dfrac{x}{8}, & 0 \le x < 4 \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de probabilidad de $Y = h(X) = 2X + 4$.

Obsérvese que $y = h(x) = 2x + 4$ es una función creciente de x. La solución inversa es $x = u(y) = (y - 4)/2$, a partir de la cual se encuentra que el jacobiano es $J = u'(y) = dx/dy = 1/2$. Por lo tanto, por el teorema la distribución de probabilidad de Y es

$$f_Y(y) = \begin{cases} \dfrac{(y-4)/2}{8}\left(\dfrac{1}{2}\right) = \dfrac{y-4}{32}, & 4 \le y \le 12 \\ 0, & \text{en caso contrario} \end{cases}$$

Se considera ahora el caso en que X_1 y X_2 son variables aleatorias continuas y quiere encontrarse la distribución de probabilidad conjunta de $Y_1 = h_1(X_1, X_2)$ y $Y_2 = h_2(X_1, X_2)$ donde la transformación es uno a uno. La aplicación de este caso será de manera típica para encontrar la distribución de probabilidad de $Y_1 = h_1(X_1, X_2)$, como en el caso de las variables discretas analizado arriba. Se necesitará el siguiente teorema.

Teorema III-4

Suponga que X_1 y X_2 son variables aleatorias **continuas** con distribución de probabilidad conjunta $f_{X_1X_2}(x_1, x_2)$, y sea que $Y_1 = h_1(X_1, X_2)$ y $Y_2 = h_2(X_1, X_2)$ definan una transformación uno a uno entre los puntos (x_1, x_2) y (y_1, y_2). Sea que las ecuaciones $y_1 = h_1(x_1, x_2)$ y $y_2 = h_2(x_1, x_2)$ tengan como soluciones únicas $x_1 = u_1(y_1, y_2)$ y $x_2 = u_2(y_1, y_2)$ para x_1 y x_2 en términos de y_1 y y_2. Entonces la probabilidad conjunta de Y_1 y Y_2 es

$$f_{Y_1Y_2}(y_1, y_2) = f_{X_1X_2}[u_1(y_1, y_2), u_2(y_1, y_2)]|J|$$

donde J es el *jacobiano*, y está dado por el determinante siguiente:

$$J = \begin{vmatrix} \partial x_1/\partial y_1, & \partial x_1/\partial y_2 \\ \partial x_2/\partial y_1, & \partial x_2/\partial y_2 \end{vmatrix}$$

y se usa el valor absoluto del determinante.

Este teorema puede usarse para encontrar $f_{Y_1Y_2}(y_1, y_2)$, la distribución de probabilidad conjunta de Y_1 y Y_2. Entonces la distribución de probabilidad de Y_1 es

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1Y_2}(y_1, y_2)dy_2$$

Es decir, $f_{Y_1}(y_1)$ es la **distribución de probabilidad marginal** de Y_1.

EJEMPLO III-4

Suponga que X_1 y X_2 son variables aleatorias exponenciales independientes con $f_{X_1}(x_1) = 2e^{-2x_1}$ y $f_{X_2}(x_2) = 2e^{-2x_2}$. Encuentre la distribución de probabilidad de $Y = X_1/X_2$.

La distribución de probabilidad conjunta de X_1 y X_2 es

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} 4e^{-2(x_1+x_2)}, & x_1 \geq 0, \quad x_2 \geq 0 \\ 0, & \text{en caso contrario} \end{cases}$$

ya que X_1 y X_2 son independientes. Sea $Y_1 = h_1(X_1, X_2) = X_1/X_2$ y $Y_2 = h_2(X_1, X_2) = X_1 + X_2$. Las soluciones inversas de $y_1 = x_1/x_2$ y $y_2 = x_1 + x_2$ son $x_1 = y_1y_2/(1 + y_1)$ y $x_2 = y_2/(1 + y_1)$, y se sigue que

$$\frac{\partial x_1}{\partial y_1} = y_2\left[\frac{1}{(1+y_1)^2}\right], \quad \frac{\partial x_1}{\partial y_2} = \left[\frac{y_1}{(1+y_1)}\right]$$

$$\frac{\partial x_2}{\partial y_1} = y_2\left[\frac{-1}{(1+y_1)^2}\right], \quad \frac{\partial x_2}{\partial y_2} = \left[\frac{1}{(1+y_1)}\right]$$

Por lo tanto,

$$J = \begin{vmatrix} \dfrac{y_2}{(1+y_1)^2} & \dfrac{y_1}{(1+y_1)} \\ \dfrac{-y_2}{(1+y_1)^2} & \dfrac{1}{(1+y_1)} \end{vmatrix} = \frac{y_2}{(1+y_1)^2}$$

y por el teorema III-4 la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$\begin{aligned} f_{Y_1Y_2}(y_1, y_2) &= f_{X_1X_2}[u_1(y_1, y_2), u_2(y_1, y_2)]|J| \\ &= 4e^{-2[y_1y_2/(1+y_1)+y_2/(1+y_1)]}\left|\frac{y_2}{(1+y_1)^2}\right| \\ &= 4e^{-2y_2}\, y_2/(1+y_1)^2 \end{aligned}$$

para $y_1 > 0$, $y_2 > 0$. Es necesario encontrar la distribución de $Y_1 = X_1/X_2$. Es la distribución de probabilidad marginal de Y_1, o

$$\begin{aligned} f_{y_1}(y_1) &= \int_0^\infty f_{y_1y_2}(y_1, y_2)\,dy_2 \\ &= \int_0^\infty 4e^{-2y_2}\left[y_2/(1+y_1)^2\right]dy_2 \\ &= \begin{cases} \dfrac{1}{(1+y_1)^2}, & y_1 > 0 \\ 0, & \text{en caso contrario} \end{cases} \end{aligned}$$

En algunos problemas donde intervienen transformaciones se requiere encontrar la distribución de probabilidad de la variable aleatoria $Y = h(X)$ cuando X es una variable aleatoria continua, pero la transformación no es uno a uno. El siguiente teorema es útil.

Teorema III-5

Suponga que X es una variable aleatoria **continua** con distribución de probabilidad $f_X(x)$, y $Y = h(X)$ es una transformación que no es uno a uno. Si puede hacerse la partición del intervalo en el que está definida X en m conjuntos, disjuntos mutuamente, excluyentes tales que cada una de las funciones inversas $x_1 = u_1(y)$, $x_2 = u_2(y)$, ..., $x_m = u_m(y)$ de $y = u(x)$ es uno a uno, entonces la distribución de probabilidad de Y es

$$f_Y(y) = \sum_{i=1}^{m} f_X[u_i(y)]|J_i|$$

donde $J_i = u_i'(y)$, $i = 1, 2, \ldots, m$ y se usa el valor absoluto.

Para ilustrar cómo se usa este teorema, suponer que X es una variable aleatoria normal con media μ y varianza σ^2, y que quiere demostrarse que la distribución de $Y = (X - \mu)^2/\sigma^2$ es una distribución ji-cuadrada con 1 grado de libertad. Sea $Z = (X - \mu)/\sigma$, y $Y = Z^2$. La distribución de probabilidad de Z es la distribución normal estándar; es decir,

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

Las soluciones inversas de $y = z^2$ son $z = \pm\sqrt{y}$, por lo que la transformación no es uno a uno. Se definen $z_1 = -\sqrt{y}$ y $z_2 = +\sqrt{y}$ de tal modo que $J_1 = -(1/2)/\sqrt{y}$ y $J_2 = (1/2)/\sqrt{y}$. Entonces por el teorema III-5, la distribución de probabilidad de Y es

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \left| \frac{-1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-y/2} \left| \frac{1}{2\sqrt{y}} \right| \\ &= \frac{1}{2^{1/2}\sqrt{\pi}} y^{1/2-1} e^{-y/2}, \quad y > 0 \end{aligned}$$

Puede demostrarse ahora que $\sqrt{\pi} = \Gamma(1/2)$, por lo que $f(y)$ puede escribirse como

$$f_Y(y) = \begin{cases} \dfrac{1}{2^{1/2}\Gamma\left(\dfrac{1}{2}\right)} y^{1/2-1} e^{-y/2}, & y > 0 \\ 0, & \text{en caso contrario} \end{cases}$$

que es la distribución ji-cuadrada con 1 grado de libertad.

EJERCICIOS

III-1. Suponga que X es una variable aleatoria con distribución de probabilidad

$$f_X(x) = \begin{cases} 1/4, & x = 1, 2, 3, 4 \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = 2X + 1$.

III-2. Sea X una variable aleatoria binomial con $p = 0.25$ y $n = 3$. Encuentre la distribución de probabilidad de la variable aleatoria $Y = X^2$.

III-3. Suponga que X es una variable aleatoria continua con distribución de probabilidad

$$f_X(x) = \begin{cases} \dfrac{x}{18}, & 0 \le x \le 6 \\ 0, & \text{en caso contrario} \end{cases}$$

a) Encuentre la distribución de probabilidad de la variable aleatoria $Y = 2X + 10$.

b) Encuentre el valor esperado de Y.

III-4. Suponga que X tiene la distribución de probabilidad uniforme

$$f_X(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{en caso contrario} \end{cases}$$

Demuestre que la distribución de probabilidad de la variable aleatoria $Y = -2\ ln\ X$ es jicuadrada con dos grados de libertad.

III-5. Una corriente de I amperes fluye a través de una resistencia de R ohms de acuerdo con la distribución de probabilidad

$$f_I(i) = \begin{cases} 2i, & 0 \le i \le 1 \\ 0, & \text{en caso contrario} \end{cases}$$

Suponga que la resistencia también es una variable aleatoria con distribución de probabilidad

$$f_R(r) = \begin{cases} 1, & 0 \le r \le 1 \\ 0, & \text{en caso contrario} \end{cases}$$

Suponga que I y R son independientes.

a) Encuentre la distribución de probabilidad de la potencia (en watts) $P = I^2R$.

b) Encuentre $E(P)$.

III-6. Una variable aleatoria X tiene la siguiente distribución de probabilidad:

$$f_X(x) = \begin{cases} e^{-x}, & x \ge 0 \\ 0, & \text{en caso contrario} \end{cases}$$

a) Encuentre la distribución de probabilidad de $Y = X^2$.

b) Encuentre la distribución de probabilidad de $Y = X^{1/2}$.

c) Encuentre la distribución de probabilidad de $Y = ln\ X$.

III-7. La velocidad de una partícula en un gas es una variable aleatoria V con distribución de probabilidad

$$f_V(v) = \begin{cases} av^2e^{-bv}, & v > 0 \\ 0, & \text{en caso contrario} \end{cases}$$

donde b es una constante que depende de la temperatura del gas y de la masa de la partícula.

a) Encuentre el valor de la constante a.

b) La energía cinética de la partícula es $W = mV^2/2$. Encuentre la distribución de probabilidad de W.

III-8. Suponga que X tiene la distribución de probabilidad

$$f_X(x) = \begin{cases} 1, & 1 \le x \le 2 \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = e^X$.

III-9. Demuestre que el teorema III-3 es válido cuando $y = h(x)$ es una función decreciente de x.

III-10. La variable aleatoria X tiene la distribución de probabilidad

$$f_X(x) = \begin{cases} \dfrac{x}{8}, & 0 \le x \le 4 \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de probabilidad de $Y = (X - 2)^2$.

III-11. Considere un rectángulo con lados de longitud S_1 y S_2, donde S_1 y S_2 son variables aleatorias independientes. Las distribuciones de probabilidad de S_1 y S_2 son

$$f_{S_1}(s_1) = \begin{cases} 2s_1, & 0 \le s_1 \le 1 \\ 0, & \text{en caso contrario} \end{cases}$$

y

$$f_{S_2}(s_2) = \begin{cases} \dfrac{s_2}{8}, & 0 \le s_2 \le 4 \\ 0, & \text{en caso contrario} \end{cases}$$

a) Encuentre la distribución conjunta del área del rectángulo $A = S_1S_2$ y la variable aleatoria $Y = S_1$.

b) Encuentre la distribución de probabilidad del área A del rectángulo.

III-12. Suponga que se tiene un circuito eléctrico simple en el que se cumple la ley de Ohm $V = IR$. Quiere encontrarse la distribución de probabilidad de la resistencia R dado que V e I son variables aleatorias independientes con las siguientes distribuciones:

$$f_V(v) = \begin{cases} e^{-v}, & v \ge 0 \\ 0, & \text{en caso contrario} \end{cases}$$

y

$$f_I(i) = \begin{cases} 1, & 1 \le i \le 2 \\ 0, & \text{en caso contrario} \end{cases}$$

Encuentre la distribución de probabilidad de R.

IV. DESARROLLO DE LAS DISTRIBUCIONES t Y F

A lo largo de este libro se ha hecho un uso extensivo de las **distribuciones de muestreo**. Dos de estas distribuciones de muestreo, las distribuciones t y F, desempeñan un papel central en la estimación estadística y en la prueba de hipótesis. Se presentan ahora las deducciones que llevan a estas distribuciones de muestreo. En esta presentación se usan las técnicas de transformación de variables de la sección III. Los resultados se enuncian como teoremas.

Teorema IV-1. La distribución t

Sea Z una variable aleatoria normal estándar y sea V una variable aleatoria ji-cuadrada con k grados de libertad. Si Z y V son independientes, entonces la distribución de la variable aleatoria

$$T = \frac{Z}{\sqrt{V/k}}$$

es la distribución t con k grados de libertad. La función de densidad de probabilidad es

$$f(t) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\,\Gamma(k/2)} \frac{1}{\left[(t^2/k)+1\right]^{(k+1)/2}}, -\infty < t < \infty$$

Demostración Puesto que Z y V son independientes, su distribución de probabilidad es

$$f_{ZV}(z,v) = \frac{v^{(k/2)-1}}{\sqrt{2\pi}\,2^{k/2}\Gamma\left(\frac{k}{2}\right)} e^{-(z^2+v)/2}, \quad -\infty < z < \infty, 0 < v < \infty$$

Aplicando el teorema III-4, se define una nueva variable aleatoria $U = V$. Por tanto, las soluciones inversas de

$$t = \frac{z}{\sqrt{v/k}}$$

y

$$u = v$$

son

$$z = t\sqrt{\frac{v}{k}}$$

y

$$v = u$$

El jacobiano es

$$J = \begin{vmatrix} \sqrt{\frac{u}{k}} & \frac{t}{2uk} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{u}{k}}$$

Por tanto,

$$|J| = \sqrt{\frac{u}{k}}$$

y la distribución de probabilidad conjunta de T y U es

$$f_{TU}(t,u) = \frac{\sqrt{u}}{\sqrt{2\pi_k}\, 2^{k/2}\Gamma\left(\frac{k}{2}\right)} u^{(k/2)-1} e^{-[(u/k)t^2+u]/2}$$

Ahora bien, puesto que $V > 0$, debe requerirse que $U > 0$, y puesto que $-\infty < Z < \infty$, entonces $-\infty < T < \infty$. Al reordenar esta última ecuación, se tiene

$$f_{TU}(t,u) = \frac{1}{\sqrt{2\pi_k}\, 2^{k/2}\Gamma\left(\frac{k}{2}\right)} u^{(k-1)/2} e^{-(u/2)[(t^2/k)+1]}, \quad 0 < \mathrm{u} < \infty, -\infty < t < \infty$$

La distribución de probabilidad de T se encuentra por medio de

$$\begin{aligned} f_T(t) &= \int_0^\infty f_{TU}(t,u)\,du \\ &= \frac{1}{\sqrt{2\pi_k}\, 2^{k/2}\Gamma\left(\frac{k}{2}\right)} \int_0^\infty u^{(k-1)/2} e^{-(u/2)[(t^2/k)+1]}\,du \\ &= \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\,\Gamma\left(\frac{k}{2}\right)} \frac{1}{[(t^2/k)+1]^{(k+1)/2}}, \qquad -\infty < t < \infty \end{aligned}$$

Esta es la distribución dada en el teorema IV-1.

La distribución de probabilidad de la variable aleatoria T fue publicada por primera vez por W.S. Gosset en su célebre escrito de 1908. Gosset era empleado de Guiness Brewers en Irlanda. Como su patrón desaprobaba la publicación de investigaciones de un empleado, Gosset publicó estos resultados bajo el pseudónimo "Student". Como resultado, esta distribución de probabilidad en ocasiones se llama **distribución *t* de Student**.

Teorema IV-2. La distribución F

Sean U_1 y U_2 variables aleatorias ji-cuadrada independientes con ν_1 y ν_2 grados de libertad, respectivamente. Entonces el cociente

$$F = \frac{U_1/\nu_1}{U_2/\nu_2}$$

tiene la función de densidad de probabilidad

$$f(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} x^{\nu_1/2-1}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)\left[\left(\frac{\nu_1}{\nu_2}\right)x+1\right]^{(\nu_1+\nu_2)/2}}, \qquad 0<x<\infty$$

Esta es la distribución F con ν_1 grados de libertad en el numerador y ν_2 grados de libertad en el denominador.

Demostración Puesto que U_1 y U_2 son variables aleatorias ji-cuadrada independientes, su distribución de probabilidad conjunta es

$$f(u_1,u_2) = \frac{u_1^{\nu_1/2-1}u_2^{\nu_2/2-1}}{2^{\nu_1/2}\Gamma\left(\frac{\nu_1}{2}\right)2^{\nu_2/2}\Gamma\left(\frac{\nu_2}{2}\right)}e^{-(u_1+u_2)/2}, \qquad 0<u_1,u_2<\infty$$

Utilizando el método del teorema III-4, se define la nueva variable aleatoria $M = U_2$. Las soluciones inversas de

$$x = \left(\frac{u_1}{\nu_1}\right)\Big/\left(\frac{u_2}{\nu_2}\right) \text{ y } m = u_2$$

son

$$u_1 = \frac{\nu_1}{\nu_2}mx \text{ y } u_2 = m$$

Por lo tanto, el jacobiano es

$$J = \begin{vmatrix} \frac{v_1}{v_2} m & \frac{v_1}{v_2} x \\ 0 & 1 \end{vmatrix} = \frac{v_1}{v_2} m$$

Por tanto, la función de densidad de probabilidad conjunta de X y M es

$$f(x,m) = \frac{\left(\frac{v_1}{v_2} mx\right)^{v_1/2-1} m^{v_2/2-1} e^{-(1/2)[(v_1/v_2)mx+m]} \left(\frac{v_1}{v_2}\right) m}{2^{v_1/2}\Gamma\left(\frac{v_1}{2}\right) 2^{v_2/2}\Gamma\left(\frac{v_2}{2}\right)}, \qquad 0 < x, m < \infty$$

La función de densidad de probabilidad de F es

$$\begin{aligned} f(x) &= \int_0^\infty f(x,m)\, dm \\ &= \frac{\left(\frac{v_1}{v_2} x\right)^{v_1/2-1} \left(\frac{v_1}{v_2}\right)}{2^{(v_1+v_2)/2}\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \int_0^\infty m^{(v_1+v_2)/2-1} e^{-(m/2)[(v_1/v_2)x+1]}\, dm \end{aligned}$$

Substituyendo $z = \frac{m}{z}\left(\frac{v_1}{v_2} x + 1\right)$ y $dm = 2\left(\frac{v_1}{v_2} x + 1\right)^{-1} dz$, obtenemos

$$\begin{aligned} f(x) &= \frac{\left(\frac{v_1}{v_2}\right)^{v_1/2} x^{v_1/2-1}}{2^{(v_1+v_2)/2}\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \int_0^\infty \left(\frac{2z}{\frac{v_1}{v_2} x + 1}\right)^{(v_1+v_2)/2-1} e^{-z} 2\left(\frac{v_1}{v_2} x + 1\right)^{-1} dz, \\ &= \frac{\left(\frac{v_1}{v_2}\right)^{v_1/2} x^{v_1/2-1}}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\left(\frac{v_1}{v_2} x + 1\right)^{(v_1+v_2)/2}} \int_0^\infty z^{(v_1+v_2)/2-1} e^{-z} dz \\ &= \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)\left(\frac{v_1}{v_2}\right)^{v_1/2} x^{v_1/2-1}}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\left(\frac{v_1}{v_2} x + 1\right)^{(v_1+v_2)/2}}, \qquad 0 < x < \infty \end{aligned}$$

que es la función de densidad de probabilidad del teorema.

V. ENFOQUE BAYESIANO DE LA ESTIMACIÓN

En este libro se utilizan métodos de inferencia estadística basados en la información de los datos muestrales. De hecho, en estos métodos las probabilidades se interpretan como frecuencias relativas. En ocasiones a las probabilidades que se interpretan de esta manera se les llama **probabilidades objetivas**. Sin embargo, hay otro enfoque para la inferencia estadística, llamado el **enfoque bayesiano**, que combina la información muestral con otra información con que se cuente antes de colectar la muestra. En esta sección se ilustra brevemente la forma en que puede usarse este enfoque en la estimación de parámetros.

Suponga que la variable aleatoria X tiene una distribución de probabilidad que es una función de un parámetro θ. Esta distribución de probabilidad se escribirá como $f(x \mid \theta)$. Esta notación implica que la forma exacta de la distribución de X es condicional respecto del valor asignado a θ. El enfoque clásico de la estimación consistiría en tomar una muestra aleatoria de tamaño n de esta distribución para sustituir después los valores muestrales x_i en el estimador de θ. Este estimador podría haberse desarrollado utilizando el enfoque de la máxima verosimilitud, por ejemplo.

Suponga que se cuenta con información adicional de θ y que dicha información puede resumirse en la forma de una distribución de probabilidad para θ, por ejemplo, $f(\theta)$. A esta distribución de probabilidad con frecuencia se le llama la **distribución previa** de θ, y suponga que la media de la distribución previa es μ_0 y que la varianza es σ_0^2. Se trata de un concepto muy novedoso en lo que al resto de este libro se refiere, porque ahora el parámetro θ se está considerando como una variable aleatoria. A las probabilidades asociadas con la distribución previa con frecuencia se les llama **probabilidades subjetivas**, por cuanto generalmente reflejan el grado de creencia del analista respecto del verdadero valor de θ. En el enfoque bayesiano de la estimación se utiliza la distribución previa de θ, $f(\theta)$, y la distribución de probabilidad conjunta de la muestra, por ejemplo $f(x_1, x_2, ..., x_n \mid \theta)$, para encontrar una **distribución posterior** de θ, por ejemplo $f(\theta \mid x_1, x_2, ..., x_n)$. Esta distribución posterior contiene información tanto de la muestra como de la distribución previa de θ. En cierto sentido, expresa nuestro grado de creencia respecto del verdadero valor de θ después de observar los datos muestrales. En el plano conceptual, es sencillo encontrar la distribución posterior. La distribución de probabilidad conjunta de la muestra $X_1, X_2, ..., X_n$ y el parámetro θ (recuérdese que θ es una variable aleatoria) es

$$f(x_1, x_2, \ldots, x_n, \theta) = f(x_1, x_2, \ldots, x_n | \theta) f(\theta)$$

y la distribución marginal de $X_1, X_2, ..., X_n$ es

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} \sum_{\theta} f(x_1, x_2, \ldots, x_n, \theta), & \theta \text{ discreta} \\ \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n, \theta)\, d\theta, & \theta \text{ continua} \end{cases}$$

Por lo tanto, la distribución deseada es

$$f(\theta | x_1, x_2, \ldots, x_n) = \frac{f(x_1, x_2, \ldots, x_n, \theta)}{f(x_1, x_2, \ldots, x_n)}$$

El **estimador de Bayes** de θ se define como el valor de $\tilde{\theta}$ que corresponde a la media de la distribución posterior $f(\theta \mid x_1, x_2, ..., x_n)$.

En ocasiones, la media de la distribución posterior de θ puede determinarse fácilmente. Como una función de θ, $f(\theta \mid x_1, x_2,..., x_n)$ es una función de densidad de probabilidad y $x_1, x_2, ..., x_n$ son tan sólo constantes. Debido a que θ entra en $f(\theta \mid x_1, x_2, ..., x_n)$ sólo a través de $f(x_1, x_2, ..., x_n, \theta)$ si $f(x_1, x_2, ..., x_n, \theta)$, como una función de θ, se identifica como una función de probabilidad conocida, entonces la media posterior de θ puede deducirse de la distribución conocida sin necesidad de integrar o incluso sin calcular $f(x_1, x_2, ..., x_n)$.

EJEMPLO V-1

Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de la distribución normal con media μ y varianza σ^2, donde no se conoce μ y σ^2 es conocida. Suponga que la distribución previa de μ es normal con media μ_0 y varianza σ_0^2 ; es decir

$$f(\mu)=\frac{1}{\sqrt{2\pi}\sigma_0}e^{-(\mu-\mu_0)^2/(2\sigma_0^2)}=\frac{1}{\sqrt{2\pi}\sigma_0^2}e^{-(\mu^2-2\mu_0\mu+\mu_0^2)/(2\sigma_0^2)}$$

La distribución de probabilidad conjunta de la muestra es

$$f(x_1,x_2,\ldots,x_n|\mu)=\frac{1}{(2\pi\sigma^2)^{n/2}}e^{-(1/2\sigma^2)\sum_{i=1}^{n}(x_i-\mu)^2}$$
$$=\frac{1}{(2\pi\sigma^2)^{n/2}}e^{-(1/2\sigma^2)(\sum x_i^2-2\mu\sum x_i-n\mu^2)}$$

Por tanto, la distribución de probabilidad conjunta de la muestra y μ es

$f(x_1, x_2, ..., x_n, \mu)$

$$=\frac{1}{(2\pi\sigma^2)^{n/2}\sqrt{2\pi}\sigma_0}e^{-(1/2)[(1/\sigma_0^2+n/\sigma^2)\mu^2-(2\mu_0/\sigma_0^2+2\sum x_i/\sigma^2)\mu+\sum x_i^2/\sigma^2+\mu_0^2/\sigma_0^2]}$$
$$=e^{-(1/2)\left[\left(\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2/n}\right)\mu^2-2\left(\frac{\mu_0}{\sigma_0^2}+\frac{\bar{x}}{\sigma^2/n}\right)\mu\right]}h_1(x_1,\ldots,x_n,\sigma^2,\mu_0,\sigma_0^2)$$

Después de completar el trinomio cuadrado perfecto en el exponente, se obtiene

$$=e^{-(1/2)\left(\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2/n}\right)\left[\mu^2-\left(\frac{(\sigma^2/n)\mu_0}{\sigma_0^2+\sigma^2/n}+\frac{\bar{x}\sigma_0^2}{\sigma_0^2+\sigma^2/n}\right)\right]^2}h_2(x_1,\ldots,x_n,\sigma^2,\mu_0,\sigma_0^2)$$

donde $h_i(x_1, x_2, ..., x_n, \sigma^2, \mu_0, \sigma_0^2)$ es una función de los valores observados, σ^2, μ_0 y σ_0^2.

Ahora bien, como $f(x_1, x_2, ..., x_n)$ no depende de μ,

$$f(\mu|x_1,\ldots,x_n)=e^{-(1/2)\left(\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2/n}\right)\left[\mu^2-\left(\frac{(\sigma^2/n)\mu_0\sigma_0^2\bar{x}}{\sigma_0^2+\sigma^2/n}\right)\right]}h_3(x_1,\ldots,x_n,\sigma^2,\mu_0,\sigma_0^2)$$

Esta expresión se identifica como una función de densidad de probabilidad normal con media posterior

$$\frac{\frac{\sigma^2}{n}\mu_0+\sigma_0^2\bar{x}}{\sigma_0^2+\frac{\sigma^2}{n}}$$

y varianza posterior

$$\left(\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2/n}\right)^{-1}=\frac{\sigma_0^2(\sigma^2/n)}{\sigma_0^2+\sigma^2/n}$$

Por consiguiente, la estimación de Bayes de μ es un promedio ponderado de μ_0 y $\bar{x}$. Para fines de comparación, obsérvese que la estimación de máxima verosimilitud de μ es $\hat{\mu}=\bar{x}$.

Para ilustrar, suponga que se tiene un tamaño de la muestra $n = 10$ de una distribución normal con media μ desconocida y varianza $\sigma^2 = 4$. Suponga asimismo que la distribución previa de μ es normal con media $\mu_0 = 0$ y varianza $\sigma_0^2 = 1$. Si la media muestral es 0.75, entonces la estimación de Bayes de μ es

$$\frac{\frac{4}{10}(0)+1(0.75)}{1+\frac{4}{10}}=\frac{0.75}{1.4}=0.536$$

Obsérvese que la estimación de máxima verosimilitud de μ es $\bar{x} = 0.75$.

Existe una relación entre el estimador de Bayes de un parámetro y el estimador de máxima verosimilitud del mismo parámetro. Para tamaños grandes de la muestra, ambos estimadores son prácticamente equivalentes. En general, la diferencia entre los dos estimadores es pequeña en comparación con $1/\sqrt{n}$. En problemas prácticos, un tamaño de la muestra moderado producirá aproximadamente la misma estimación ya sea con el método de Bayes que con el de máxima verosimilitud, si los resultados muestrales son consistentes con la información previa supuesta. Si los resultados muestrales no son consistentes con los supuestos previos, entonces la estimación de Bayes puede diferir considerablemente de la estimación de máxima verosimilitud. Bajo estas circunstancias, si se acepta que los resultados muestrales son correctos, la información previa debe ser incorrecta. Entonces sería mejor utilizar la estimación de máxima verosimilitud.

Si los resultados muestrales son muy diferentes de la información previa, el estimador de Bayes tenderá siempre a producir una estimación que está entre la estimación de máxima verosimilitud y los supuestos previos. Si hay una inconsistencia mayor entre la información previa y la muestra, la diferencia entre las dos estimaciones será más grande.

Los métodos bayesianos pueden usarse para construir estimaciones por intervalos de parámetros que son similares a los intervalos de confianza. Si se ha obtenido la densidad poste-

rior de θ, puede construirse un intervalo, por lo general con centro en la media posterior, que contiene $100(1-\alpha)$ por ciento de la probabilidad posterior. A tal intervalo se le llama **intervalo de Bayes** del $100(1-\alpha)$ por ciento para el parámetro desconocido θ.

Mientras que en muchos casos la estimación del intervalo de Bayes para θ será muy similar a un intervalo de confianza típico con el mismo coeficiente de confianza, la interpretación de los dos es muy diferente. Un intervalo de confianza es un intervalo que, antes de que se tome la muestra, incluirá al parámetro desconocido θ con una probabilidad de $1-\alpha$. Es decir, el intervalo de confianza típico se relaciona con la frecuencia relativa de un intervalo que incluye a θ. Por otra parte, un intervalo de Bayes es un intervalo que contiene $100(1-\alpha)$ por ciento de la probabilidad posterior de θ. Puesto que la densidad de probabilidad posterior mide un grado de creencia acerca de θ dados los resultados muestrales, el intervalo de Bayes proporciona un grado de creencia subjetivo acerca de θ en lugar de una interpretación de frecuencia. La estimación del intervalo de Bayes de θ es afectado por los resultados muestrales pero no está determinado por completo por ellos.

EJEMPLO V-2

Suponga que la variable aleatoria X tiene una distribución normal con media μ y varianza 4. No se conoce el valor de μ, pero una distribución previa razonable sería la distribución normal con media 2 y varianza 1. Puede demostrarse que la distribución de probabilidad posterior de μ es una distribución normal con media $[(n/4)+1]^{-1}[(n\bar{x}/4)+2]$ y varianza $[(n/4)+1]^{-1}$. Por lo tanto, un intervalo de Bayes del $100(1-\alpha)$ por ciento para m que es simétrico alrededor de la media posterior es

$$[(n/4)+1]^{-1}[(n\bar{x}/4)+2]-z_{\alpha/2}[(n/4)+1]^{-1/2} \le \mu \le [(n/4)+1]^{-1}[(n\bar{x}/4)+2]+z_{\alpha/2}[(n/4)+1]^{-1/2}$$

Suponga ahora que se toma una muestra aleatoria de tamaño $n = 16$ y que la media muestral es $\bar{x} = 2.5$. Un intervalo de Bayes de 95% para μ sería

$$1.52 \le \mu \le 3.48$$

El intervalo de Bayes es ligeramente más corto que el intervalo de confianza típico, debido a que usar la información previa es equivalente a un ligero incremento del tamaño de la muestra.

EJERCICIOS

V-1. Suponga que X es una variable aleatoria normal con media m desconocida y varianza σ^2 conocida. La distribución previa de μ es una distribución normal con media μ_0 y varianza σ_0^2.
a) Encuentre la distribución posterior de μ.
b) Encuentre el estimador de Bayes de μ.

V-2. Suponga que X es una variable aleatoria normal con media μ desconocida y varianza σ^2 conocida. La distribución previa de μ es una distribución uniforme definida en el intervalo $[a, b]$.
a) Encuentre la distribución posterior de μ.
b) Encuentre el estimador de Bayes de μ.

V-3. Suponga que X es una variable aleatoria de Poisson con parámetro λ. Sea la distribución previa de λ una distribución gama con parámetros $m+1$ y $(m+1)/\lambda_0$.
a) Encuentre la distribución posterior de λ.

b) Encuentre el estimador de Bayes de λ.

V-4. Suponga que X es una variable aleatoria normal con media desconocida y varianza $\sigma^2 = 9$ conocida. La distribución previa de μ es normal con media $\mu_0 = 4$ y varianza $\sigma_0^2 = 1$. Se toma una muestra aleatoria de $n = 25$ observaciones, y la media muestral es $\bar{x} = 4.85$.

a) Encuentre el estimador de Bayes de μ.

b) Establezca una estimación del intervalo de Bayes de 95% para μ.

c) Compare el intervalo de Bayes del inciso b) con un intervalo de confianza típico de 95% para μ. ¿Cuál de los intervalos es más largo?

V-5. El peso de unas cajas de dulces es una variable aleatoria normal con media μ y varianza 1/10 lb. La distribución previa de μ es normal con media 5.03 lb y varianza 1/25 lb. Una muestra aleatoria de 10 cajas da una media muestral de $\bar{x} = 5.05$ lb.

a) Encuentre la estimación de Bayes de μ.

b) Encuentre un intervalo de Bayes de 95% para μ.

c) Si las cajas que pesan menos de 4.95 lb se consideran defectuosas, ¿qué porcentaje de las cajas producidas en este proceso es defectuoso?

V-6. El tiempo entre las fallas de una máquina tiene una distribución exponencial con parámetro λ. Suponga que la distribución previa de λ es exponencial con una media de 100 horas. Se observan dos máquinas, y el tiempo promedio entre las fallas es $\bar{x} = 1\,125$ horas.

a) Encuentre la estimación de Bayes de λ.

b) ¿Qué proporción de las máquinas se esperaría que falle antes de 1 000 horas?

VI. PRUEBAS DEL COCIENTE DE VEROSIMILITUD

La prueba de hipótesis es una de las técnicas más importantes de la inferencia estadística. A lo largo de este libro se han presentado muchas aplicaciones de la prueba de hipótesis. Aun cuando se ha hecho hincapié en un desarrollo heurístico, muchos de estos procedimientos de prueba de hipótesis pueden desarrollarse utilizando un principio general llamado el principo del cociente de verosimilitud. Las pruebas desarrolladas con este método con frecuencia resultan "mejores" procedimientos de prueba en el sentido de que minimizan la probabilidad β del error tipo II entre todas las pruebas que tienen la misma probabilidad α del error tipo I.

Es sencillo ilustrar el principio del cociente de verosimilitud. Suponga que la variable aleatoria X tiene una distribución de probabilidad que está descrita por un parámetro desconocido θ, por ejemplo, $f(x, \theta)$. Quiere probarse la hipótesis H_0: θ está en Ω_0 contra H_1: θ está en Ω_1, donde Ω_0 y Ω_1 son conjuntos disjuntos de valores (tales como H_0: $\mu \geq 0$ contra H_1: $\mu < 0$). Sean $x_1, x_2, \ldots, x_n$ las observaciones de una muestra aleatoria. La distribución conjunta de estas observaciones muestrales es

$$f(x_1, x_2, \ldots, x_n, \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \cdots \cdot f(x_n, \theta)$$

Recuérdese por la revisión de la estimación de máxima verosimilitud del capítulo 7 que la función de verosimilitud, por ejemplo $L(\theta)$, es tan sólo esta distribución conjunta considerada como una función del parámetro θ. El **principio del cociente de verosimilitud** para la construcción de pruebas consta de los pasos siguientes:

1. Encuentre el valor más grande de la verosimilitud para cualquier θ en Ω_0. Esto se hace encontrando el estimador de máxima verosimilitud de θ restringido a los valores

incluidos en Ω_0 y sustituyendo este valor de θ de nuevo en la función de verosimilitud. Se obtiene así un valor de la función de verosimilitud al que se llamará $L(\Omega_0)$.

2. Encuentre el valor más grande de la verosimilitud para cualquier θ en Ω_1. A este valor de la función de verosimilitud se le llama $L(\Omega_1)$.

3. Forme el cociente

$$\lambda = \frac{L(\Omega_0)}{L(\Omega_1)}$$

A este cociente λ se le llama el **estadístico de la prueba del cociente de verosimilitud**.

Este procedimiento de prueba requiere rechazar la hipótesis nula H_0 cuando el valor de este cociente λ es pequeño, por ejemplo, siempre que $\lambda < k$, donde k es una constante. Por tanto, el principio del cociente de verosimilitud requiere rechazar H_0 cuando $L(\Omega_1)$ es mucho mayor que $L(\Omega_0)$, lo cual indicaría que los mismos datos son más compatibles con la hipótesis alternativa H_1 que con la hipótesis nula H_0. Generalmente, la constante k se seleccionaría para obtener un valor especificado de α, la probabilidad del error tipo I.

Estas ideas pueden ilustrarse con un problema de prueba de hipótesis que ya se ha estudiado: el de probar si la media de una población normal tiene un valor especificado μ_0. Suponga que se tiene una muestra de n observaciones de una población normal con media μ desconocida y varianza σ^2 desconocida, por ejemplo, $x_1, x_2, \ldots, x_n$. Quiere probarse la hipótesis H_0: $\mu = \mu_0$ contra H_1: $\mu \neq \mu_0$. La función de verosimilitud de la muestra es

$$L = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^{n} (x_i-\mu)^2/(2\sigma^2)}$$

y los valores de Ω_0 y Ω_1 son $\Omega_0 = \mu_0$ y $\Omega_1 = \{\mu: -\infty < \mu < \infty\}$, respectivamente. Los valores de μ y σ^2 que maximizan L en Ω_1 son las estimaciones de máxima verosimilitud usuales de μ y σ^2:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

Al sustituir estos valores en L, se tiene

$$L(\Omega_1) = \left[\frac{1}{(2\pi/n)\sum(x_i - \bar{x})^2}\right]^{n/2} e^{-(n/2)}$$

Para maximizar L en Ω_0 tan sólo se hace $\mu = \mu_0$ y después se encuentra el valor de σ^2 que maximiza L. Se encuentra que este valor es

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu_0)^2$$

de donde se obtiene

$$L(\Omega_0) = \left[\frac{1}{(2\pi/n)\sum(x_i - \mu_0)^2}\right]^{n/2} e^{-(n/2)}$$

El cociente de verosimilitud es

$$\lambda = \frac{L(\Omega_0)}{L(\Omega_1)} = \left[\frac{\sum(x_i - \bar{x})^2}{\sum(x_i - \mu_0)^2}\right]^{n/2}$$

Ahora bien, puesto que

$$\sum (x_i - \mu_0)^2 = \sum(x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2$$

λ puede escribirse como

$$\lambda = \left\{\frac{1}{1 + \dfrac{n(\bar{x} - \mu_0)^2}{\sum(x_i - \bar{x})^2}}\right\}^{n/2}$$

$$= \left\{\frac{1}{1 + \left[\dfrac{n(\bar{x} - \mu_0)^2}{\sum(x_i - \bar{x})^2}\right]\left(\dfrac{n-1}{n-1}\right)}\right\}^{n/2}$$

$$= \left\{\frac{1}{1 + \left(\dfrac{1}{n-1}\right)\left[\dfrac{(\bar{x} - \mu_0)^2}{s^2/n}\right]}\right\}^{n/2}$$

Obsérvese que $\left[\dfrac{(\bar{x} - \mu_0)^2}{s^2/n}\right] = t^2$ es el cuadrado del valor de una variable aleatoria que tiene la distribución t con $n - 1$ grados de libertad cuando la hipótesis nula H_0: $\mu = \mu_0$ es verdadera. Por tanto, el valor del cociente de verosimilitud λ puede escribirse como

$$\lambda = \left[\frac{1}{1+\left[t^2/(n-1)\right]}\right]^{n/2}$$

Es sencillo encontrar el valor de la constante k que llevaría a rechazar la hipótesis nula H_0. Puesto que H_0 se rechaza si $\lambda < k$, esto implica que los valores pequeños de λ apoyan la hipótesis alternativa. Evidentemente, λ será pequeña cuando t^2 es grande. Por tanto, en vez de especificar k puede especificarse una constante c y H_0: $\mu = \mu_0$ se rechaza si $t^2 > c$. Los valores críticos de c serían los valores extremos, sean positivos o negativos, y si quiere controlarse la probabilidad del error tipo I en α, la región crítica en términos de t sería

$$t < -t_{\alpha/2,\, n-1} \text{ y } t > t_{\alpha/2,\, n-1}$$

o, de manera equivalente, se rechazaría H_0: $\mu = \mu_0$ si $t^2 > c = t^2_{\alpha/2,\, n-1}$. Por lo tanto, la prueba del cociente de verosimilitud para H_0: $\mu = \mu_0$ contra H_1: $\mu \neq \mu_0$ es la familiar prueba de una sola muestra.

El procedimiento empleado en este ejemplo para encontrar la región crítica para el cociente de verosimilitud λ es de uso común. Es decir, de manera típica, λ puede manipularse para producir una condición que sea equivalente a $\lambda < k$, pero una condición que es más fácil de usar.

El principio del cociente de verosimilitud es un procedimiento muy general. La mayoría de las pruebas presentadas en este libro que utilizan las distribuciones t, ji-cuadrada y F para probar medias y varianzas de distribuciones normales son pruebas del cociente de verosimilitud. El principio también puede usarse en casos en que las observaciones son dependientes, o incluso en casos en que sus distribuciones son diferentes. Sin embargo, la función de verosimilitud puede ser muy complicada en algunas de estas situaciones. Para usar el principio del cociente de verosimilitud debe especificarse la forma de la distribución. Sin dicha especificación, es imposible escribir la función de verosimilitud y, en consecuencia, si uno no está dispuesto a suponer una distribución de probabilidad particular, el principio del cociente de verosimilitud no puede usarse. Esto podría llevar al uso de los procedimientos de prueba no paramétricos comentados en el capítulo 14.

VII. FACTORES ALEATORIOS EN EXPERIMENTOS FACTORIALES

El capítulo 13 sobre experimentos factoriales se enfocó principalmente en el caso en que todos los factores eran fijos; es decir, el experimentador elige de manera específica los niveles, y las conclusiones del experimento se restringen a dichos niveles específicos. Se considerará ahora brevemente la situación en que uno o más de los factores de un experimento factorial son aleatorios, utilizando el diseño factorial de dos factores como ilustración.

El modelo de efectos aleatorios

Suponga que se tienen dos factores A y B dispuestos en un experimento factorial en el que los niveles de ambos factores se seleccionan al azar de poblaciones grandes de niveles de los factores, y que las conclusiones quieren ampliarse a la población completa de los niveles de los factores. Las observaciones se representan con el modelo

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{VII-1}$$

donde los parámetros τ_i, β_j, $(\tau\beta)_{ij}$ y ε_{ijk}, son variables aleatorias que tienen distribuciones normales e independientes con medias cero y varianzas $\sigma_\tau^2, \sigma_\beta^2, \sigma_{\tau\beta}^2$, y σ^2, respectivamente. Como resultado de estos supuestos, la varianza de cualquier observación Y_{ijk} es

$$V(Y_{ijk}) = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2$$

y a $\sigma_\tau^2, \sigma_\beta^2, \sigma_{\tau\beta}^2$, y σ^2 se les llama los **componentes de la varianza**. Las hipótesis que quieren probarse son H_0: $\sigma_\tau^2 = 0$, H_0: $\sigma_\beta^2 = 0$, y H_0: $\sigma_{\tau\beta}^2 = 0$. Obsérvese la similitud con el modelo de efectos aleatorios del experimento con un solo factor revisado en el capítulo 12.

El análisis de varianza básico se mantiene sin cambios; es decir, SS_A, SS_B, SS_{AB}, SS_T y SS_E se calculan como en el caso de los efectos fijos. Para construir el estadístico de la prueba, deben examinarse los cuadrados medios esperados. Éstos son

$$E(MS_A) = \sigma^2 + n\sigma_{\tau\beta}^2 + n\sigma_\beta^2$$
$$E(MS_B) = \sigma^2 + n\sigma_{\tau\beta}^2 + an\sigma_\beta^2$$
$$E(MS_{AB}) = \sigma^2 + n\sigma_{\tau\beta}^2$$

y

$$E(MS_E) = \sigma^2 \tag{VII-2}$$

Obsérvese a partir de los cuadrados medios esperados que el estadístico apropiado para probar la hipótesis de no interacción H_0: $\sigma_{\tau\beta}^2 = 0$ es

$$F_0 = \frac{MS_{AB}}{MS_E} \tag{VII-3}$$

ya que si H_0 es verdadera, tanto el numerador como el denominador de F_0 tienen como esperado a σ^2, y únicamente si H_0 es falsa $E(MS_{AB})$ es mayor que $E(MS_E)$. El cociente F_0 se distribuye como $F_{(a-1)(b-1),\, ab(n-1)}$. De manera similar, para probar que no hay ningún efecto principal del factor A, o H_0: $\sigma_\tau^2 = 0$, se usaría

$$F_0 = \frac{MS_A}{MS_{AB}} \tag{VII-4}$$

que se distribuye como $F_{a-1,(a-1)(b-1)}$, y para probar H_0: $\sigma_\beta^2 = 0$ el estadístico de la prueba es

$$F_0 = \frac{MS_B}{MS_{AB}} \tag{VII-5}$$

que se distribuye como $F_{b-1,(a-1)(b-1)}$. Todas estas pruebas son de cola superior. Por tanto, las hipótesis nulas anteriores deberán rechazarse con el nivel de significación α si el valor calculado de f_0 excede el punto porcentual α superior de la distribución F. Obsérvese que estos estadísticos de prueba no son los mismos que se usaron cuando los dos factores A y B son fijos. Siempre se usan los cuadrados medios esperados como guía para construir el estadístico de prueba.

Los componentes de la varianza pueden estimarse igualando los cuadrados medios observados con sus valores esperados y despejando los componentes de la varianza. Se obtiene así

$$\begin{aligned} \hat{\sigma}^2 &= MS_E \\ \hat{\sigma}_{\tau\beta}^2 &= \frac{MS_{AB} - MS_E}{n} \\ \hat{\sigma}_\beta^2 &= \frac{MS_B - MS_{AB}}{an} \\ \hat{\sigma}_\tau^2 &= \frac{MS_A - MS_{AB}}{bn} \end{aligned} \tag{VII-6}$$

EJEMPLO VII-1

Se estudian dos factores que pueden influir en la resistencia a la ruptura de una tela. Se eligen al azar cuatro máquinas y tres operadores de prueba, y se hace un experimento utilizando tela del mismo segmento de producción. Los datos se presentan en la tabla VII-1, y el análisis de varianza se muestra en la tabla VII-2. Obsérvese que las cuatro primeras columnas de la tabla VII-2 se calculan como en el análisis estándar (modelo de efectos fijos). Los estadísticos de la prueba se calculan con las ecuaciones (VII-3) a (VII-5). Se usará $\alpha = 0.05$. El estadístico de la prueba para la hipótesis de no interacción H_0: $\sigma_{\tau\beta}^2 = 0$ es

$$f_0 = \frac{MS_{AB}}{MS_E} = \frac{5.94}{3.75} = 1.584$$

Tabla VII-1 Datos de la resistencia a la ruptura del ejemplo VII-1

	Máquina			
Operador	1	2	3	4
A	113	113	111	113
	112	118	111	119
B	111	110	111	114
	112	111	109	112
C	109	112	114	111
	111	115	112	112

el cual se compararía con $f_{0.05,\,6,\,12} = 3.00$, y por tanto H_0: $\sigma^2_{\tau\beta} = 0$ no puede rechazarse. El valor P para este cociente es $P = 0.2338$. Para probar que no hay ningún efecto de la máquina (H_0: $\sigma^2_\beta = 0$), se calcula

$$f_0 = \frac{MS_B}{MS_{AB}} = \frac{8.11}{5.94} = 1.365$$

que se compararía con $f_{0.05,\,3,\,6} = 4.76$. El valor P para este cociente es $P = 0.3403$. Por lo tanto, se concluye que las máquinas no afectan significativamente los resultados de la prueba de la resistencia a la ruptura. Para probar que no hay ningún efecto del operador (H_0: $\sigma^2_\tau = 0$), se calcula

$$f_0 = \frac{MS_A}{MS_{AB}} = \frac{13.17}{5.94} = 2.217$$

que se compararía con $f_{0.05,\,2,\,6} = 5.14$. El valor P para este cociente es $P = 0.1904$. Se concluye que los operadores no afectan los resultados de la prueba de la resistencia a la ruptura. Los componentes de la varianza pueden estimarse utilizando las ecuaciones VII-6 de la siguiente manera:

Tabla VII-2 Análisis de varianza del ejemplo VII-1

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f_0	Valor P
Operadores	26.33	2	13.17	2.217	0.1904
Máquinas	24.33	3	8.11	1.365	0.3403
Interacción	35.67	6	5.94	1.584	0.2338
Error	45.00	12	3.75		
Total	131.33	23			

$$\hat{\sigma}^2 = 3.75$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{5.94 - 3.75}{2} = 1.10$$

$$\hat{\sigma}^2_{\tau} = \frac{13.17 - 5.94}{8} = 0.90$$

$$\hat{\sigma}^2_{\beta} = \frac{8.11 - 5.94}{6} = 0.36$$

Todos los componentes de la varianza $\hat{\sigma}^2_{\tau}$, $\hat{\sigma}^2_{\beta}$ y $\hat{\sigma}^2_{\tau\beta}$ son pequeños; de hecho, no son significativamente diferentes de cero.

El modelo mixto

Suponga ahora que uno de los factores, A, es fijo y el otro, B, es aleatorio. A esta situación se le llama el análisis de varianza de **modelo mixto**. El modelo lineal es

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \qquad \text{(VII-7)}$$

En este modelo τ_i es un efecto fijo definido de tal modo que $\sum_{i=1}^{a} \tau_i = 0$, β_j es un efecto aleatorio, el término de interacción $(\tau\beta)_{ij}$ es un efecto aleatorio, y ε_{ijk} es una variable aleatoria que tiene una distribución normal e independiente con media cero y varianza σ^2. Los elementos de interacción $(\tau\beta)_{ij}$ son variables aleatorias normales con media cero y varianza $[(a-1)/a]\,\hat{\sigma}^2_{\tau\beta}$, y $\sum_{i=1}^{a} (\tau\beta)_{ij} = 0$. Como la suma de los efectos de interacción en los niveles del factor fijo es igual a cero, a esta versión del modelo mixto se le llama el **modelo restringido**. Los elementos de interacción no son todos independientes. Para mayores detalles, véase Montgomery.

Los cuadrados medios esperados en este caso son

$$\begin{aligned} E(MS_A) &= \sigma^2 + n\sigma^2_{\tau\beta} + \frac{bn\sum_{i=1}^{a} \tau_i^2}{a-1} \\ E(MS_B) &= \sigma^2 + an\sigma^2_{\beta} \\ E(MS_{AB}) &= \sigma^2 + n\sigma^2_{\tau\beta} \\ E(MS_E) &= \sigma^2 \end{aligned} \qquad \text{(VII-8)}$$

Por lo tanto, el estadístico de la prueba apropiado para probar H_0: $\tau_i = 0$ es

$$F_0 = \frac{MS_A}{MS_{AB}} \tag{VII-9}$$

para el que la distribución de referencia es $F_{a-1,(a-1)(b-1)}$. Para probar H_0: $\sigma_\beta^2 = 0$, el estadístico de la prueba es

$$F_0 = \frac{MS_B}{MS_E} \tag{VII-10}$$

para el que la distribución de referencia es $F_{b-1,ab(n-1)}$. Por último, para probar H_0: $\sigma_{\tau\beta}^2 = 0$ se usaría

$$F_0 = \frac{MS_{AB}}{MS_E} \tag{VII-11}$$

para el que la distribución de referencia es $F_{(a-1)(b-1),ab(n-1)}$.

Los componentes de la varianza σ_β^2, $\sigma_{\tau\beta}^2$ y σ^2 pueden estimarse eliminando la primera ecuación de las ecuaciones VII-8, dejando tres ecuaciones con tres incógnitas, cuya solución es

$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_E}{an}$$

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n}$$

y

$$\hat{\sigma}^2 = MS_E$$

Este enfoque general puede usarse para estimar los componentes de la varianza de *cualquier* modelo mixto. Después de eliminar los cuadrados medios que contienen factores fijos, quedará

Tabla VII-3 Análisis de varianza del modelo mixto con dos factores (A fijo, B aleatorio)

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio esperado	Cuadrado medio	F_0
Renglones (A)	SS_A	$a-1$	$\sigma^2 + n\sigma^2_{\tau\beta} + bn\sum\tau_i^2/(a-1)$	MS_A	MS_A/MS_{AB}
Columnas (B)	SS_B	$b-1$	$\sigma^2 + an\sigma^2_{\beta}$	MS_B	MS_B/MS_E
Interacción	SS_{AB}	$(a-1)(b-1)$	$\sigma^2 + n\sigma^2_{\tau\beta}$	MS_{AB}	MS_{AB}/MS_E
Error	SS_E	$ab(n-1)$	σ_2	MS_E	
Total	SS_T	$abn-1$			

siempre un conjunto de ecuaciones que pueden resolverse para los componentes de la varianza. En la tabla VII-3 se resume el análisis de varianza del modelo mixto con dos factores.

Salida de computadora

Algunos paquetes de software tienen la capacidad para manejar factores aleatorios en el análisis de varianza (SAS, Minitab, Statgraphics, por ejemplo). La tabla VII-4 es la solución de computadora del ejemplo VII-1 de Minitab. Obsérvese que las pruebas F de los efectos principales se han construido utilizando MS_{AB} en el denominador. Los resultados concuerdan bastante bien con los de la tabla VII-2.

La tabla VII-5 es la salida de Minitab para los datos de la resistencia a la ruptura suponiendo un modelo mixto, donde los operadores son fijos y las máquinas son aleatorias. Minitab supone que se ha especificado la forma restringida del modelo mixto [no todos los paquetes de software trabajan con este supuesto: SAS usa un modelo no restringido, es decir, el término $\sum_{i=1}^{a}(\tau\beta)_{ij} \neq 0$ en general]. Obsérvese que los estadísticos de las pruebas para los efectos principales y las interacciones se han construido apropiadamente utilizando las ecuaciones VII-9 a VII-11.

Tabla VII-4 Salida de Minitab para el ejemplo VII-1

Análisis de varianza (diseños balanceados)

Factor	Tipo	Niveles	Valores			
Operador	aleatoria	3	A	B	C	
Máquina	aleatoria	4	1	2	3	4

Análisis de varianza de la resistencia

Fuente	GL	SS	MS	F	P
Operador	2	26.333	13.167	2.21	0.190
Máquina	3	24.333	8.111	1.36	0.340
Operador*Máquina	6	35.667	5.944	1.59	0.234
Error	12	45.000	3.750		
Total	23	131.333			

Tabla VII-5 Análisis de varianza de Minitab para los datos de la resistencia a la ruptura de la tabla VII-1 donde los operadores son fijos y las máquinas son aleatorias (modelo mixto)

Análisis de varianza (diseños balanceados)

Factor	Tipo	Niveles	Valores			
Operador	fija	3	A	B	C	
Máquina	aleatoria	4	1	2	3	4

Análisis de varianza para la resistencia

Fuente	GL	SS	MS	F	P
Operador	2	26.333	13.167	2.21	0.190
Máquina	3	24.333	8.111	2.16	0.145
Operador*Máquina	6	35.667	5.944	1.59	0.234
Error	12	45.000	3.750		
Total	23	131.333			

EJERCICIOS

VII-1. Considérese el experimento descrito en el ejemplo 13-4. Suponga que ambos factores son aleatorios.

a) Analice los datos y saque las conclusiones apropiadas.

b) Estime los componentes de la varianza.

VII-2. Para los datos de la resistencia a la ruptura de la tabla VII-1, suponga que los operadores se eligieron al azar, pero las máquinas fueron un factor fijo. ¿Afecta esto el análisis o las conclusiones?

VII-3. Una compañía emplea dos ingenieros de programación de operaciones. Su supervisora quiere determinar si los estándares establecidos por ellos están influidos por una interacción entre ingenieros y operadores. Ella selecciona tres operadores al azar y realiza un experimento en el que los ingenieros establecen tiempos estándares para la misma tarea. Ella obtiene los datos que se muestran abajo.

	Operador		
Ingeniero	1	2	3
1	2.59	2.38	2.40
	2.78	2.49	2.72
2	2.15	2.85	2.66
	2.86	2.72	2.87

a) Enuncie las hipótesis apropiadas.

b) Use el análisis de varianza para probar estas hipótesis con $\alpha = 0.05$.

c) Analice gráficamente los residuales de este experimento.

d) Estime los componentes de la varianza apropiados.

VII-4. Considere el experimento de la densidad del endurecimiento de un ánodo de carbono descrito en el ejercicio 13-4. Suponga que las posiciones en el horno se eligieron al azar y la temperatura es un factor fijo.

a) Enuncie las hipótesis apropiadas.

b) Use el análisis de varianza para probar estas hipótesis con $\alpha = 0.05$.

c) Estime los componentes de la varianza.

VII-5. Considere el experimento descrito en el ejercicio VII-3. ¿Cómo cambia el análisis (y las conclusiones) si ambos factores son aleatorios? Use $\alpha = 0.05$.

Apéndice C

Bibliografía

OBRAS DE INTRODUCCIÓN Y MÉTODOS GRÁFICOS

Chambers, J., W. Cleveland, B. Kleiner y P. Tukey, *Graphical Methods for Data Analysis*, Wadsworth & Brooks/Cole, Pacific Grove, CA. Una presentación muy bien redactada de los métodos gráficos en la estadística.

Freedman, D., R. Pisani, R. Purves y A. Adbikari, *Statistics*, 2a. edición, Norton, Nueva York. Una excelente introducción al razonamiento estadístico que requiere antecedentes matemáticos mínimos.

Hoaglin, D., F. Mosteller y J. Tukey, *Understanding Robust and Exploratory Data Analysis*, John Wiley & Sons, Nueva York. Un buen análisis y demostración de técnicas como las representaciones de tallo y hoja y los diagramas de caja.

Tanur, J., *et al.* (eds.), *Statistics: A Guide to the Unknown*, 3a. edición, Wadsworth & Brooks/Cole, Pacific Grove, CA. Contiene una colección de artículos no matemáticos breves que describen diferentes aplicaciones de la estadística.

Tukey, J., *Exploratory Data Analysis*, Addison-Wesley, Reading, MA. Introduce varios métodos descriptivos y analíticos nuevos. No muy fácil de leer.

PROBABILIDAD

Derman, C., L. Gleser e I. Olkin, *Probability Models and Applications*, Macmillan, Nueva York. Un tratamiento completo de la teoría de la probabilidad con un nivel matemático más elevado que el de este libro.

Hoel, P.G., S.C. Port y C.J. Stone, *Introduction to Probability Theory*, Hougthon Mifflin, Boston. Un tratamiento completo y bien escrito de la teoría de la probabilidad y las distribuciones continuas estándar de variables discretas y continuas.

Mosteller, F., R. Rourke y G. Thomas, *Probability with Statistical Applications*, 2a. ed., Addison-Wesley, Reading, MA. Una introducción, sin cálculo, a la probabilidad con muchos ejemplos excelentes.

Ross, S., *A First Course in Probability*, 3a. ed., Macmillan, Nueva York. Con un nivel matemático más elevado que este libro, pero incluye muchos ejemplos y ejercicios excelentes.

ESTADÍSTICA MATEMÁTICA

Efron, B. y R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, Nueva York. Una referencia importante sobre esta técnica útil pero con uso intensivo de computadora.

Hoel, P.G., *Introduction to Mathematical Statistics*, 5a. ed., John Wiley & Sons, Nueva York. Notable libro de introducción, bien escrito y en general de fácil comprensión.

Hogg, R. y A. Craig, *Introduction to Mathematical Statistics*, 4a. ed., Macmillan, Nueva York. Otra obra clásica sobre los principios matemáticos de la estadística; de un nivel más elevado que el libro de Hoel, pero contiene excelentes revisiones de la estimación y la prueba de hipótesis.

Larsen, R. y M. Marx, *Introduction to Mathematical Statistics*, 2a. ed., Prentice-Hall, Englewood Cliffs, NJ. Escrito en un nivel matemático relativamente bajo, de fácil lectura.

Larson, H.J., *Introduction to Probability Theory and Statistical Inference*, 3a. ed., John Wiley & Sons, Nueva York. Un libro muy bien escrito que ofrece una amplia cobertura de muchos aspectos de la probabilidad y la estadística matemática.

ESTADÍSTICA PARA INGENIERÍA

Devore, J.L., *Probability and Statistics for Engineering and the Sciences*, 4a. ed., Wadsworth & Brooks/Cole, Pacific Grove, CA. Trata muchos de los tópicos del presente libro, pero con un nivel matemático ligeramente más elevado. Muchos de los ejemplos y ejercicios involucran aplicaciones a las ciencias biológicas y de la vida.

Hines, W.W. y D.C. Montgomery, *Probability and Statistics in Engineering and Management Science*, 3a. ed., John Wiley & Sons, Nueva York. Trata muchos de los tópicos del presente libro. Hace mayor énfasis en la probabilidad y maneja un nivel matemático más elevado.

Ross, S., *Introduction to Probability and Statistics for Engineers and Scientists*, John Wiley & Sons, Nueva York. Escrito con un estilo y una orientación matemática más rigurosa que este libro, pero contiene buenos ejemplos.

Walpole, R.E., R.H. Myers y S.L. Myers, *Probability and Statistics for Engineers and Scientists*, 6a. ed., Prentice-Hall, Inc., Upper Saddle River, Nueva Jersey. Un libro muy bien escrito más o menos en el mismo nivel que éste.

ANÁLISIS DE REGRESIÓN

Daniel, C. y F. Wood, *Fitting Equations to Data*, 2a. ed., John Wiley & Sons, Nueva York. Excelente referencia que contiene muchos detalles del análisis de datos.

Draper, N. y H. Smith, 3a. ed., John Wiley & Sons, Nueva York. *Applied Regression Analysis*, 2a. ed., John Wiley & Sons, Nueva York. Un libro de gran alcance sobre la regresión escrito para lectores con una orientación estadística.

Montgomery, D.C. y E.A. Peck, *Introduction to Linear Regression Analysis*, 2a. ed., John Wiley & Sons, Nueva York. Un detallado libro sobre la regresión escrito para ingenieros y estudiantes de ciencias físicas.

Myers, R.H., *Classical and Modern Regression with Applications*, 2a. ed., PWS-Kent, Boston. Contiene muchos ejemplos con la salida del programa SAS anotada. Muy bien escrito.

Neter, J., W. Wasserman, C. Nachtsheim y M. Kutner, *Applied Linear Statistical Models*, 4a. ed., Richard D. Irwin, Homewood, Ill. La primera parte del libro es una introducción a la regresión lineal simple y múltiple. Se orienta a la administración y la economía.

Younger, M.S., *A Handbook for Linear Regression*, 2a. ed. Duxburg, Boston. Una buena presentación de los métodos de regresión. La revisión de los paquetes de computadora SAS, BMD y SPSS es excelente.

DISEÑO DE EXPERIMENTOS

Box, G.E.P., W.G. Hunter y J.S. Hunter, *Statistics for Experimenters*, John Wiley & Sons, Nueva York. Una excelente introducción para aquellos lectores que desean un tratamiento con una orientación estadística. Contiene muchas sugerencias útiles para el análisis de datos.

Mason, R.L., R.F. Gunst y J.F. Hess, *Statistical Design and Analysis of Experiments*, John Wiley & Sons, Nueva York. Un libro detallado que trata la estadística básica, la prueba de hipótesis y los intervalos de confianza, los aspectos elementales del diseño experimental y el análisis de regresión.

Montgomery, D.C., *Diseño y Análisis de Experimentos*, 2a. ed., Limusa-Wiley. Escrito en el mismo nivel que el libro de Box, Hunter y Hunter, pero enfocado en las aplicaciones de ingeniería.

ESTADÍSTICA NO PARAMÉTRICA

Conover, W.J., *Practical Nonparametric Statistics*, 2a. ed., John Wiley & Sons, Nueva York. Una excelente exposición de los métodos de la estadística no paramétrica, incluye muchos buenos ejemplos y ejercicios.

Hollander, M. y D. Wolfe, *Nonparametric Statistical Methods*, John Wiley & Sons, Nueva York. Un buen libro de referencia, con una serie de tablas muy útiles.

CONTROL ESTADÍSTICO DE CALIDAD Y MÉTODOS RELACIONADOS

Duncan, A.J., *Quality Control and Industrial Statistics*, 5a. ed., Richard D. Irwin, Homewood, Illinois. Un libro clásico sobre el tema.

Grant, E.L. y R.S. Leavenworth, *Statistical Quality Control*, 6a. ed., McGraw-Hill, Nueva York. Uno de los primeros libros sobre el tema; contiene muchos buenos ejemplos.

John, P.W.M., *Statistical Methods in Engineering and Quality Improvement*, John Wiley & Sons, Nueva York. No es un libro de métodos, sino una presentación bien escrita de la metodología estadística para el mejoramiento de calidad.

Montgomery, D.C., *Introduction to Statistical Quality Control*, 3a. ed., John Wiley & Sons, Nueva York. Un tratamiento moderno del tema escrito en el mismo nivel que este libro.

Nelson, W., *Applied Life Data Analysis*, John Wiley & Sons, Nueva York. Contiene muchos ejemplos del uso de los métodos estadísticos para el estudio de datos de fallas; una buena referencia para los aspectos estadísticos de la ingeniería de confiabilidad y las distribuciones de probabilidad especiales usadas en ese campo.

Ryan, T.P., *Statistical Methods for Quality Improvement*, John Wiley & Sons, Nueva York. Ofrece un tratamiento amplio del tema, con cierto énfasis en las técnicas más recientes.

Wadsworth, H.M., K.S. Stephens y A.B. Godfrey, *Modern Methods for Quality Control and Improvement*, John Wiley & Sons, Nueva York. Un tratamiento extenso de los métodos estadísticos para el mejoramiento de calidad con un nivel un tanto más elevado que este libro.

Western Electric Company, *Statistical Quality Control Handbook*, Western Electric Company, Inc., Indianapolis, Indiana. Algo viejo pero bueno.

Apéndice D

Respuestas de ejercicios seleccionados

CAPÍTULO 1

Sección 1-1 y 1-2

1-1. 74.004, 0.000022414, 0.00473

1-3. 7068.1, 51302.265, 226.5

1-5. 43.975, 151.143, 12.294

1-7. Ejercicio 1-4

Ejercicios complementarios

1-9. **a)** 2.17 **b)** 0.43, 0.656

1-13. **a)** 83.11 **b)** 50.6, 7.114

1-19. $\bar{y}$ = 446.11°C, s_y = 5.833°C

CAPÍTULO 2

Sección 2-2

2-5.

Variable	N	Mediana	Q1	Q3
octanaje	82	90.400	88.575	92.209

2-7.

Variable	N	Mediana	Q1	Q3
algodón	64	34.798	34.700	34.7

Sección 2-4

2-15. **a)** Cuartil superior: 75
Cuartil inferior: 60
b) Mediana: 67.5
c) Cuartil superior: 75
Cuartil inferior: 58.5
Mediana: 68

2-17. **a)** Media muestral: 2.415
Desviación estándar muestral: 0.534

2-19. **a)** Cuartil superior: 84.5
Cuartil inferior: 79.5
c) Cuartil superior: 83.75
Cuartil inferior: 79.25

Ejercicios complementarios

2-27. **a)** 7.1838
b) s^2 = 0.000427, s = 0.02066

2-29. **a)** Rango muestral = 3.2
s^2 = 0.8663, s = 0.9308
b) Rango muestral = 3.2
s^2 = 0.8663, s = 0.9308

2-31. **c)** Primeras 40 observaciones:
Media muestral = 14.875
Varianza muestral = 0.89936

Segundas 40 observaciones:
Media muestral = 14.9225
Varianza muestral = 1.0464

2-33. **b)** 9.325, 4.4858

CAPÍTULO 3

Sección 3-1

3-1. Sea que "a" y "b" denoten una pieza arriba y abajo, respectivamente, de las especificaciones.
$S = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$

3-3. Sea que "a" denote una fuente de poder satisfactoria.
Sea que "f", "m" y "c" denoten una fuente con un error funcional, secundario y de acabado, respectivamente.
$S = \{a, f, m, c\}$

3-5. $S = \{0, 1, 2, ..., 24\}$, el espacio muestral contiene 2^{24} vectores.

3-7. S es el espacio muestral de 100 posibles enteros de dos dígitos.

3-9. Si P es el conjunto de los enteros no negativos, entonces $S = P \times P \times P$.

3-15. $S = \{s, fs, ffs, fffS, fffFS, fffFFS, fffFFFA\}$

3-21. Sea que "d" denote una calculadora defectuosa y sea que "o" denote una calculadora satisfactoria.
a) $S = \{ddd, odd, ddo, odo, dod, ood, doo, ooo\}$
b) $A = \{ddd, ddo, dod, doo\}$
c) $B = \{ddd, ddo, odd, odo\}$
d) $A \cap B = \{ddd, ddo\}$
e) $B \cup C = \{ddd, ddo, odd, odo, dod, ood\}$

3-23.

	Número de muestras
$A' \cap B$	10
B'	15
$A \cup B$	92

3-25. **a)** $A' = \{x \mid x \geq 72.5\}$
b) $B' = \{x \mid x \leq 52.5\}$
c) $A \cap B = \{x \mid 52.5 < x < 72.5\}$
d) $A \cup B = \{x \mid x > 0\}$

Sección 3-2

3-27. Todos los resultados son igualmente probables.
a) 2/5 **b)** 3/5 **c)** 3/5 **d)** 1 **e)** 0

3-29. **a)** $S = \{1, 2, 3, 4, 5, 6\}$ **b)** 1/6
c) 2/6 **d)** 5/6

3-31. 6

3-33. **a)** 1/10 **b)** 5/10

3-35. **a)** 86/100 **b)** 89/100 **c)** 14/100
d) 80/100 **e)** 0.95 **f)** 0.94

3-37. **a)** 20/40 **b)** 35/40 **c)** 20/40
d) 18/40 **c)** 37/40 **f)** 38/40

Sección 3-3

3-39. **a)** 0.7 **b)** 0.4 **c)** 0.1 **d)** 0.2 **e)** 0.6
f) 0.8

3-41. Si A, B y C son mutuamente excluyentes, entonces $P(A \cup B \cup C) = P(A) + P(B) + P(C) = 0.3 + 0.4 + 0.5 = 12$, que es mayor que 1. Por lo tanto $P(A)$, $P(B)$ y $P(C)$ no pueden ser iguales a los valores dados.

3-43. **a)** 350/370 **b)** 362/370 **c)** 358/370
d) 345/370

Sección 3-4

3-45. **a)** 86/100 **b)** 89/100 **c)** 80/89
d) 80/86

3-47. **a)** 345/357 **b)** 5/13

3-49. **a)** 0.15 **b)** 11/72 **c)** 0.72 **d)** 11/15
e) 0.11 **f)** 0.76

3-51. **a)** 15/40 **b)** 14/39 **c)** 0.135
d) = 0.599 **e)** 0.046 **f)** 0.089

3-53. **a)** 0.0060 **b)** 0.0080
c) 4.82×10^{-7}

Sección 3-5

3-57. 0.22

3-59. 0.023

3-61. 0.028

3-63. **a)** 0.20
b) Sea que C denote el evento de que la segunda pieza seleccionada tenga una contracción excesiva.
c) 0.20

Sección 3-6

3-65. No, no son independientes.

3-67. No, no son independientes.

3-69. No, no son independientes.

3-71. **a)** Sí, son independientes.
b) Sí, son independientes.

3-73. **a)** 0.000976 **b)** 0.000976 **c)** 0.246

3-75. 0.934

3-77. 0.000008

Sección 3-7

3-79. 0.08

3-81. **a)** 0.25 **b)** 0.95 c. 0.995 **d)** 0.2375
e) 0.74625 **f)** 0.98 **g)** 0.98375
h) 0.769

Sección 3-8

3-83. Continua

3-85. Continua

3-87. Discreta

3-89. Discreta

3-91. Continua

Ejercicios complementarios

3-93. **a)** 0.10 **b)** 5/8 **c)** 0.9

3-95. **a)** 0.282 **b)** 0.718

3-97. **a)** 0.0037 **b)** 0.1892

3-99. **a)** 7.59×10^{-5} **b)** 0.0022
c) 0.0244

3-101. **a)** $n = 3$ **b)** $n = 3$

3-103. **a)** 0.00285 **b)** 0.687

CAPÍTULO 4

Sección 4-1

4-1. $\{0, 1, 2, \ldots, 1000\}$

4-3. $\{0, 1, 2, \ldots, 99999\}$

4-5. $\{0, 1, \ldots, 491\}$

4-7. $\{0, 1, \ldots.\}$

4-9. $\{0, 1, \ldots, 15\}$

Sección 4-2

4-11. $f(0) = 1/3, f(1.5) = 1/3, f(2) = 1/6,$
$f(3) = 1/6$

4-13. Todas las probabilidades son mayores que 0 y menores que 1 y suman uno.
a) 1 **b)** 0 **c)** 3/4 **d)** 1/2

4-15. Todas las probabilidades son mayores que 0 y menores que 1 y suman uno.
a) 9/25 **b)** 6/25 **c)** 12/25 **d)** 0

4-17. $P(X = 10 \text{ millones}) = 0.3,$
$P(X = 5 \text{ millones}) = 0.6,$
$P(X = 1 \text{ millón}) = 0.1$

4-19. **a)** 0.512
b) $P(X = 0) = 0.008,$
$P(X = 1) = 0.096,$
$P(X = 2) = 0.384,$
$P(X = 3) = 0.512$

4-21. **a)** 0.92169
b) $P(X = 0) = 0.00001,$
$P(X = 1) = 0.00167,$
$P(X = 2) = 0.07663,$
$P(X = 3) = 0.92169$

Sección 4-3

4-23. $F(x)=\begin{cases}0, & x<0\\ 1/3, & 0\le x<1.5\\ 2/3, & 1.5\le x<2\\ 1 & 2\le x\end{cases}$

4-27. $F(x)=\begin{cases}0, & x<1\\ 4/7, & 1\le x<2\\ 6/7, & 2\le x<3\\ 1 & 3\le x\end{cases}$

4-29. **a)** 1 **b)** 0.75 **c)** 0.25 **d)** 0.25 **e)** 0 **f)** 0

Sección 4-4

4-31. $E(X)=2,\ \ V(X)=2$

4-33. $E(X)=5.30,\ \ V(X)=0.0168$

4-35. $E(X)=6.1$ millones, $V(X)=7.89$ (millones)^2

4-37. $x=30$

Sección 4-5

4-39. $E(X)=2,\ \ V(X)=3/4$

4-41. $E(X)=0.17,\ \ V(X)=0.0002$

4-43. $E(Y)=5E(X),\ \ E(Y)=22.5,$ $V(Y)=25,\ \ V(X)=205.25,$ $\sigma(X)=14.36$

Sección 4-6

4-47. **a)** 0.246 **b)** 0.055 **c)** 0.011 **d)** 0.322

4-49. **a)** 2.40×10^{-8} **b)** 0.9999 **c)** 9.91×10^{-18} **d)** 0.322

4-53. **a)** 0.368 **b)** 0.632 **c)** 0.920 **d)** $E(X)=1,\ V(X)=0.999$

4-55. **a)** $n=50, p=0.1$ **b)** 0.112 **c)** 4.51×10^{-48}

4-57. **a)** 0.996 **b)** 0.989

4-59. **a)** 0 **b)** 0.2137

Sección 4-7

4-61. **a)** 0.5 **b)** 0.0625 **c)** 0.0039 **d)** 0.75 **e)** 0.25

4-63. **a)** 0.0064 **b)** 0.9984 **c)** 0.008

4-65. **a)** 0.1024 **b)** 0.1074

4-67. **a)** $E(X)=5$ **b)** $E(X)=5$

4-69. **a)** 20 **b)** 0.0436 **c)** 0.0459 **d)** 0.0411 **e)** 19

4-71. **a)** 3 000 **b)** 1 731.18

Sección 4-8

4-75. **a)** 0.4623 **b)** 0.00021 **c)** 0.1644

4-79. **a)** 0.0424 **b)** 0.4571

4-81. **a)** 8.15×10^{-8} **b)** 2.05×10^{-5} **c)** 0.00105 **d)** 12, 271, 512

4-83. **a)** 0.5016 **b)** 0.4984 **c)** 0.3583 **d)** 2/3

Sección 4-9

4-85. **a)** 0.6703 **b)** 0.9921 **c)** 0.000715 **d)** 1.09×10^{-8}

4-87. **a)** 0.0378 **b)** 0.0103 **c)** 0.0516 **d)** 0.1755

4-89. **a)** 4.54×10^{-5} **b)** 0.6321

4-91. **a)** 0.8521 **b)** 0.3812

Ejercicios complementarios

4-93. **a)** 0.1877 **b)** 0.4148 **c)** 15

4-95. **a)** 0.011 **b)** 8/3

4-97. 0.9810

4-99. **a)** 100 **b)** 99.5

4-101. 0.1

4-103. **a)** 0.6 **b)** 0.8 **c)** 0.7 **d)** 3.9 **e)** 3.09

4-105. **a)** 0.2408 **b)** 0.4913

4-107. **a)** 0.3233 **b)** 0.092
4-109. 0.63 segundos

CAPÍTULO 5

Sección 5-2

5-1. **a)** 0.3679 **b)** 0.2858 **c)** 0.9817 **d)** 0.0498
5-3. **a)** 1 **b)** 0.6321 **c)** 3679 **d)** 0.018 **e)** 6.303
5-5. **a)** 0.05 **b)** 0.233 **c)** 0.632 **d)** 105.6
5-7. **a)** 0.25 **b)** 0.5 **c)** 0.75

Sección 5-3

5-11. **a)** 0.95 **b)** 1/8 **c)** 0 **d)** 1/2

Sección 5-4

5-21. $E(X) = 8/3$, $V(X) = 4/3$
5-23. $E(X) = 50$, $V(X) = 0.0208$
5-25. *2*
5-27. **a)** $E(X) = 1\,205$, $V(X) = 8.333$, $\sigma(X) = 2.887$

Sección 5-5

5-29. **a)** $E(X) = 0$, $V(X) = 1/3$, $\sigma(X) = 0.577$ **b)** 0.9
5-31. **b)** 0.3 **c)** 0.96 **d)** $E(X) = 1$, $V(X) = 0.00083$
5-33. **a)** 0.5 **b)** 31 **c)** $E(X) = 35$, $V(X) = 8.33$

Sección 5-6

5-35. **a)** 0.68268 **b)** 0.9545 **c)** 0.9973 **d)** 0.00135 **e)** 0.34134
5-37. **a)** 1.96 **b)** 2.58 **c)** 1 **d)** 3
5-39. **a)** 10 **b)** 0.95 **c)** 8.96 **d)** 3.92 **e)** 5.16
5-41. **a)** 5 **b)** −1.56 **c)** 6.44 **d)** No tiene solución
5-43. **a)** 0.99379 **b)** 0.00621
5-45. **a)** 0 **b)** 0.0241 **c)** (12.142, 12.658)
5-47. **a)** 0.02275 **b)** 0.47725 **c)** 0.336
5-49. **a)** 0.9733 **b)** 9.973
5-51. **a)** 0.00043 **b)** 6 016 **c)** 1/8
5-53. **a)** 0.02775 **b)** 0.324 **c)** 11.455

Sección 5-8

5-59. **a)** 0.0078 **b)** 0.023 **c)** 0.497
5-61. **a)** 0.0718 **b)** 0.047 **c)** 0.0038, 0.0062
5-63. 0.5
5-65. **a)** 0.951 **b)** 0.921 **c)** 0.966 **d)** 0.24
5-67. **a)** 0.543 **b)** 0.421

Sección 5-9

5-69. **a)** 0 **b)** 0.0183 **c)** 0.8647 **d)** 0.117 **e)** 0.0256
5-71. **a)** 0.3679 **b)** 0.2835 **c)** 0.117
5-73. **a)** 0.1353 **b)** 0.5134 **c)** 0.2031 **d)** 34.54 minutos
5-75. **a)** 0.3679 **b)** 0.3679 **c)** 2 horas
5-77. **a)** 23.03 **b)** 23.03 **c)** 6.93
5-79. **a)** 0.0409 **b)** 0.1353 **c)** 0.1353
5-81. **a)** 0.9179 **b)** 0.2336
5-83. 0.3528
5-85. **a)** 0.3528 **b)** 0.04979 **c)** 46.05
5-87. **a)** 0.3679 **b)** 0.1353 **c)** 0.0498
5-89. **a)** 0.0819 **b)** 4.95

Sección 5-10

5-91. **b)** $E(X) = 2$, $V(X) = 0.4$ **c)** 10

5-93. **a)** 1 500 **b)** 387.3 libras

5-95. **a)** 5×10^5 **b)** 223, 607 bits **c)** 0.0803

5-97. **a)** 120 días **b)** 0.4335

Sección 5-11

5-103. **a)** 0.9189 **b)** 0.1123

5-105. **a)** 0.2750 **b)** 0.6731

5-107. Exponencial $E(X) = 1\ 000$

Ejercicios complementarios

5-111. **a)** 0.3935 **b)** 0.3834 **c)** 23.03

5-113. **a)** 3.77×10^{-6} **b)** 50

5-115. 0.9933

5-117. **a)** 0.23 **b)** 0.683 **c)** 1.86

5-119. **a)** 0.006 **b)** 0.012 **c)** 5.33

5-121. (0.0008, 0.0032)

5-123. 11 158

CAPÍTULO 6

Sección 6-1

6-3. $E(X) = 2.1875$, $E(Y) = 3.625$

6-5. 1/36

6-7. $E(X) = 2.167$, $V(X) = 0.639$, $E(Y) = 2.167$, $V(Y) = 0.639$

6-13. $X \geq 0, Y \geq 0, X + Y \leq 4$

6-15. **c)** 0.0879 **d)** 0.4322

6-17. No, no son independientes.

6-19. No, no son independientes.

Sección 6-2

6-21. **a)** 0.5 **b)** 0.35 **c)** 0.5 **d)** 0.7 **e)** 1.5

6-25. **a)** 0.0289 **b)** 0.0866 **c)** 0.4989

6-29. **a)** 0.4762 **b)** 0.1333

6-33. X es binomial con $n = 3$ y $p = 0.01$, $E(X) = 0.03$, $V(X) = 0.0297$.

Sección 6-3

6-35. **a)** 0.6944 **b)** 0.6944 **c)** 0.5833 **d)** 0.3733 **e)** 2 **f)** 2

6-37. 2/39

6-41. **a)** 1/81 **b)** 5/54 **c)** 1/18 **d)** 16/81 **e)** 6/5 **f)** 4/5

6-43. 10

6-47. **a)** 0.99326 **b)** 0.00669 **c)** 0.00619 **d)** 0.99387 **e)** 1/5 **f)** 8/15

6-49. 1/7.5

6-53. Normal, $\mu = 2$ y $\sigma = 0.6$

Sección 6-4

6-55. **a)** 0.25 **b)** 0.0625 **c)** 1 **d)** 1 **e)** 2/3

6-57. **a)** $2x$ **b)** 0.25

6-59. **a)** 0.7887 **b)** 1/4

6-61. **a)** 0.4215 **b)** 0.8190

Sección 6-5

6-63. $\sigma_{XY} = 0.5313$, $\rho_{XY} = 0.937$

6-65. $c = 1/36$ $\sigma_{XY} = -0.027$, $\rho_{XY} = -0.042$

6-69. $c = 1/7.5$, $\sigma_{XY} = 1.1437$, $\rho_{XY} = 0.87891$

6-71. $\sigma_{XY} = 0$, $\rho_{XY} = 0$

Sección 6-6

6-75. 0.622

Sección 6-7

6-79. **a)** 30 **b)** 97 **c)** 0.5 **d)** 0.846

6-81. **a)** $E(T) = 3$, $V(T) = 0.1414$
b) 0.017

6-83. $D = A - B - C$
a) $E(D) = 6$ mm, $\sigma_T = 0.1225$
b) 0.206

6-85. **a)** 0.002 **b)** $n = 6$

Sección 6-8

6-89. 0.0498, 0.0183

6-91. **a)** 8×10^{-5} **b)** 3.75×10^{-5}

Ejercicios complementarios

6-93. **a)** 1/8 **b)** 1/2 **c)** 1/2 **d)** 3/8
6-97. –0.641
6-103. **a)** 1/108 **b)** 0.5787 **c)** 3/4
d) 0.2199 **e)** 9/4 **f)** 4/3

6-105. **a)** 1/2 **b)** 1/4

6-107. 3/4

6-109. **a)** 0.085
b) Z es binomial con $n = 10$, $p = 0.3$
c) 3

6-111. 0.499

6-113. **a)** 0.057 **b)** 0.057

CAPÍTULO 7

Sección 7-3

7-1. Estimador 1

7-3. Estimador 2

7-5. 2.5

7-7. El estimador 3 es más eficiente, el estimador 2 es mejor estimador.

7-11. $\hat{\lambda} = 6.2$

Sección 7-6

7-23. 0.8504

7-25 Se reduce en 0.929

7-27. 22.361

7-31. **a)** 1 **b)** 1 **c)** 0

7-33. 0.221

Ejercicios complementarios

7-39. 0.4726

7-41. No, el teorema del límite central establece que con muestras grandes, la media muestral tiene una distribución aproximadamente normal.

CAPÍTULO 8

Sección 8-1

8-1. **a)** 0.02275 **b)** 0.1587

8-3. $c = 11.48$

8-5. **a)** 0.0244 **b)** 0.0122 **c)** 0

8-7. **a)** 0.0571 **b)** 0.057

8-9. **a)** 0.0228 **b)** 0.022

8-11. **a)** 0.0892 **b)** 0.2866

8-13. (4.827, 5.173)

8-17. **a)** 0.149 **b)** 0.624 **c)** 0.376

Sección 8-2

8-19. **a)** No se rechaza H_0 **b)** 0.99865
c) 0 **d)** $96.7 \leq \mu \leq 99.31$

8-21. **a)** No se rechaza H_0 **b)** 0.15562
c) 60 **d)** 0.84134
e) $1.496 \leq \mu \leq 1.513$

8-23. **a)** Se rechaza H_0 **b)** 0.00539 **c)** 0
d) 9 **e)** $1\,003 \leq \mu \leq 1\,025$ **f)** $1\,005 \leq \mu$

8-25. 37

8-27. 30

Sección 8-3

8-29. $t_0 = 4.47$, se rechaza H_0

8-31. **a)** $t_0 = -0.154$, no se rechaza H_0
b) Valor $P > 0.80$ **c)** $n = 10$
d) $16.455 \leq \mu \leq 17.505$

8-33. **b)** $t_0 = 4.65$, se rechaza H_0
c) Valor $P > 0.0005$
d) $8.216 \leq \mu \leq 8.244$

8-35. **b)** -1.748, no se rechaza H_0
c) $2.811 \leq \mu \leq 2.99$

8-37. $230.21 \leq \mu \leq 233.13$

Sección 8-4

8-39. **a)** 11.52, no se rechaza H_0
b) $0.2 <$ Valor $P < 1.0$
c) $3.3 \leq \sigma \leq 8.76$

8-41. **a)** 118.75, se rechaza H_0
b) Valor $P < 0.01$
c) $0.031 \leq \sigma^2 \leq 0.176$

8-43. 0.6

Sección 8-5

8-45. $0.799 \leq p \leq 0.847$

8-47. **a)** $0.23 \leq p \leq 0.49$ **b)** 2 213
c) 2 401

8-49. $p \leq 0.0217$

8-51. **a)** $t_0 = -0.51$, no se rechaza H_0
b) 0.59612

8-53. **a)** $t_0 = -11.352$, se rechaza H_0 **b)** 0

8-55. **a)** $t_0 = -0.943$, se rechaza H_0
b) 0.82639

8-57. $t_0 = 1.583$, no se rechaza H_0

Sección 8-7

8-59. **a)** 9.655, no se rechaza H_0 **b)** 0.2237

8-61. **a)** 10.39, se rechaza H_0 **b)** 0.0155

8-63. **a)** 769.57, se rechaza H_0 **b)** 0

Sección 8-8

8-65. No se rechaza H_0, valor $P = 0.85$

8-67. No se rechaza H_0, valor $P = 0.652$

8-69. Se rechaza H_0, valor $P \approx 0$

Ejercicios complementarios

8-75. **a)** 40 **b)** 23

8-77. **a)** 0.083 **b)** 0.066 **c)** 0.059
e) 0.109, 0.086, 0.077

8-79. **a)** 0.64058 **b)** 0.28434 **c)** 0.0537

8-81. **a)** $n = 25$: 0.3783;
$n = 100$: 0.2643;
$n = 400$: 0.1056;
$n = 2\,500$: 0.0009
b) $n = 25$: 0.3783;
$n = 100$: 0.2643;
$n = 400$: 0.1056;
$n = 2\,500$: 0.0009

8-83. **a)** $n = 100$: $1 - \beta = 0.51595$;
$n = 150$: $1 - \beta = 0.69146$;
$n = 300$: $1 - \beta = 0.93822$;
b) $n = 100$: $1 - \beta = 0.2776$;
$n = 150$: $1 - \beta = 0.44828$;
$n = 300$: 0.81859
c) 0 **d)** 49

8-85. **a)** Se rechaza H_0 **b)** Se rechaza H_0

8-89. **a)** (0.00045, 0.0095) **b)** $n = 516$
c) $n = 25\,900$

8-91. **a)** Se rechaza H_0 **b)** 0.0123

CAPÍTULO 9

Sección 9-2

9-1. **a)** No se rechaza H_0 **b)** 0.322 **c)** 1
d) (–0.0098, 0.0298) **e)** 3

9-3. $n \cong 11$ (usando la ecuación 9-6);
$\beta \approx 0$, Potencia = 1, por la curva OC

9-5. **a)** $0.0987 \leq \mu_1 - \mu_2 \leq 0.2813$
b) $0.81 \leq \mu_1 \leq \mu_2 \leq 0.299$
c) $\mu_1 - \mu_2 \leq 0.2813$

9-7. **a)** $-3.684 \leq \mu_1 - \mu_2 \leq -2.116$
b) Se rechaza H_0 **c)** 0

9-9. 11

9-11. **a)** $-5.83 \leq \mu_1 - \mu_2 \leq -0.57$
b) Sí

9-13. Se rechaza H_0, valor $P = 0.0171$

Sección 9-3

9-15. **a)** No se rechaza H_0
b) Valor $P > 0.80$
c) $-0.415 \leq \mu_1 - \mu_2 \leq 0.515$

9-17. **a)** Se rechaza H_0
b) $-5.480 \leq \mu_1 - \mu_2 \leq 0.519$

9-19. **a)** Se rechaza H_0
b) $0.010 <$ Valor $P < 0.020$
c) $-0.749 \leq \mu_1 - \mu_2 \leq -0.115$

9-21. **a)** $-44.76 \leq \mu_1 - \mu_2 \leq -17.24$

9-23. Se rechaza H_0, valor $P < 0.0010$

9-25. **a)** Se rechaza H_0
b) $0.005 <$ valor $P < 0.010$
c) Se rechaza H_0

Sección 9-4

9-29. $0.1694 \leq \mu_d \leq 0.3778$

9-31. No se rechaza H_0

9-33. $-727.46 \leq \mu_d \leq 2\,464.21$

9-35. $t_0 = 5.645$, se rechaza H_0

9-37. $t_0 = 8.387$, se rechaza H_0

9-39. $t_0 = 3.45$, se rechaza H_0

Sección 9-5

9-41. **a)** 1.59 **b)** 2.28 **c)** 2.64 **d)** 0.529
e) 0.524 **f)** 0.311

9-43. $f_0 = 0.657$, no se rechaza H_0

9-45. $f_0 = 1.232$, no se rechaza H_0

9-47. **a)** $0.08775 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 3.594$
b) $0.00585 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 5.4$ **c)** $0.137 \leq \frac{\sigma_1^2}{\sigma_2^2}$

9-49. $f_0 = 0.297$, no se rechaza H_0

9-51. $f_0 = 0.2575$, se rechaza H_0

9-53. $0.3369 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 2.640$

Sección 9-6

9-55. **a)** 1.49, no se rechaza H_0 **b)** 0.13622

9-57. **a)** 0.81859 **b)** 383

9-59. $-0.0074 \leq p_1 - p_2 \leq 0.054$

Ejercicios complementarios

9-61. $f_0 = 0.061$, se rechaza H_0

9-63. **a)** (0.1583, 5.157)

9-65. **a)** $-5.362 \leq \mu_1 - \mu_2 \leq -0.638$
b) $-6.340 \leq \mu_1 - \mu_2 \leq 0.340$

9-67. **a)** $Z_0 = 6.55$, se rechaza H_0 **b)** Se rechaza H_0

9-69. **a)** $t_0 = 0.715$, no se rechaza H_0
b) No se rechaza H_0

9-71. **a)** $z_0 = -5.36$, se rechaza H_0

9-77. $n = 16$, sí

9-79. $n = 3$, sí

9-81 **b)** Valor $P > 0.80$ **c)** $n = 20$

CAPÍTULO 10

Sección 10-2

10-1. **a)** $\hat{\beta}_0 = 48.013$ $\hat{\beta}_1 = -2.33$
b) 37.99 **c)** 39.39 **d)** 6.71

10-3. **a)** $\hat{\beta}_0 = 0.4631476$, $\hat{\beta}_1 = 0.0074902$
b) $\hat{\beta}_1 = 0.00749$

10-5. **a)** $\hat{\beta}_0 = 13.3202$, $\hat{\beta}_1 = 3.32437$
b) 38.253 **c)** –2.02732

10-7. **a)** $\hat{\beta}_0 = 33.5348$, $\hat{\beta}_1 = 0.0353971$
b) 28.26 **c)** 1.50048

10-9. **b)** $\hat{\beta}_0 = -9.8131$, $\hat{\beta}_1 = 0.171484$
c) 4.76301

10-11. **b)** $\hat{\beta}_0 = 0.470467$, $\hat{\beta}_1 = 20.5673$
c) 21.038 **d)** 1.6629

10-17. $\hat{\beta}_0 = 0$, $\hat{\beta}_1 = 21.031461$

Sección 10-5

10-19. **a)** $f_0 = 73.95$, se rechaza H_0 **b)** 0.00796
c) se($\hat{\beta}_0$) = 0.04095,
se($\hat{\beta}_1$) = 0.000484

10-21. **a)** $t_0 = 8.518$, se rechaza H_0
b) $f_0 = 72.5563$, se rechaza H_0
c) 8.76775, se($\hat{\beta}_0$) = 2.57172,
se($\hat{\beta}_1$) = 0.390276
d) $t_0 = 5.1795$, se rechaza H_0

10-23. **a)** $f_0 = 4.53$, no se rechaza H_0
b) 3.6596 , se($\hat{\beta}_0$) = 2.614,
se($\hat{\beta}_1$) = 0.0166281
c) $t_0 = 0.87803$, no se rechaza H_0
d) $t_0 = 12.83$, se rechaza H_0

10-25. **a)** $f_0 = 44.6567$, se rechaza H_0
Valor $P = 0.000003$
b) 1.98
c) se($\hat{\beta}_0$) = 2.1356,
se($\hat{\beta}_1$) = 0.0256613
d) $t_0 = -4.5957$, se rechaza H_0,
Valor $P = 0.00022$

10-27. **a)** $f_0 = 155$, se rechaza H_0
b) 9 811.21, se ($\hat{\beta}_0$) = 2.9668,
se ($\hat{\beta}_1$) = 45.3468
c) $t_0 = -2.3466$, no se rechaza H_0
Valor $P = 0.0306$
d) $t_0 = 57.8957$, se rechaza H_0
e) $t_0 = 2.765$, no se rechaza H_0

Sección 10-6 y 10-7

10-31. **a)** (–2.917478, –1.7421254)
b) (46.714, 49.311)
c) (41.284, 43.093)
d) (39.095, 45.2826)

10-33. **a)** (–0.00961, –0.00444)
b) (16.2448, 27.3318)
c) (7.914, 10.3717)
d) (4.072, 14.214)

10-35. **a)** (9.101, 9.315)
b) (–11.622, –1.049)
c) (498.72, 501.53)
d) (495.57, 504.67)

10-37. **a)** (0.03689, 0.10183)
b) (44.0897, 49.1185)
c) (44.0897, 49.1185)
d) (37.8298, 55.3784)

10-39. **a)** (201.552, 266.59)
b) (–4.67, –2.347)
c) (111.834, 145.79)

10-41. **a)** (–43.1964, –30.7272)
b) (2 530.09, 2 720.68)
c) (1 823.783, 1 948.525)
d) (9.8201, 32.2559)

10-43. **d)** $R^2 \equiv 76.73\%$

10-45. **a)** $R^2 = 82.9\%$ **c)** Sí

10.47. **a)** $R^2 = 71.27\%$
b) $f_0 = 0.196$, no se rechaza H_0

10.49. **a)** $R^2 = 85.22\%$ **b)** $SS_{PE} = 4.505$
c) $f_0 = 8.864$, no se rechaza H_0

Sección 10-10

10-55. **a)** $\hat{\beta}_0 = -0.0280411$, $\hat{\beta}_1 = 0.990987$
b) $f_0 = 4.41$, se rechaza H_0 **c)** 0.903
d) $t_0 = 8.9345$, se rechaza H_0
e) $z_0 = 3.879$, se rechaza H_0
f) (0.7677, 0.9615)

10-57. **a)** –0.738
b) Se rechaza H_0, valor $P = 0.00000738$
c) (–0.871, –0.504)
d) No se rechaza H_0, valor $P = 0.6936$

10-59. **a)** Se rechaza H_0, valor $P = 0$
b) (0.3358, 0.8007) **c)** Sí

10-61. **a)** 0.933 **b)** $t_0 = 10.06$, se rechaza H_0
c) $\hat{\beta}_0 = 0.72538$, $\hat{\beta}_1 = 0.498081$, $f_0 = 101.16$, se rechaza H_0
d) $t_0 = 0.468345$, no se rechaza H_0

Ejercicios complementarios

10-65. **a)** $\hat{\beta}_0 = 93.34$, $\hat{\beta}_1 = 15.64$
b) $f_0 = 4.60$, se rechaza H_0
c) (7.961, 23.322)
d) (74.758, 111.923)
e) (26.18, 138.70)
f) $f_0 = 2.002$, no se rechaza H_0

10-67. **b)** $\hat{\beta}_0 = -0.8819$, $\hat{\beta}_1 = 0.00385$
c) $f_0 = 122.03$, se rechaza H_0 **d)** No
e) $\hat{\beta}_0 = 0.5967$, $\hat{\beta}_1 = 0.00097$, sí

10-69. $\hat{\beta}_1 = 0.7916$

CAPÍTULO 11

Secciones 11-2 y 11-3

11-1. **b)** $\hat{\beta} = \begin{pmatrix} 171.054 \\ 3.713 \\ -1.126 \end{pmatrix}$ **c)** 189.48

11-5. **a)** $\hat{y} = 33.4491 - 0.05435x_1 + 1.07822x_2$
b) 21.457 mpg

11-7. **a)** $\hat{y} = 350.9943 - 1.272x_1 - 0.1539x_2$
b) 165.29
c) $\hat{y}' = 125.8655 - 7.7586x_1 - 0.0943x_2 - 0.0092x_1x_2$
d) 184.13

11-9. **a)** $\hat{y} = 47.174 - 9.7352x_1 + 0.4283x_2 + 18.2375x_3$
b) 91.43

11-11. **a)** $\hat{y} = -8.0119 + 0.494x_1 + 0.0018x_2 + 0.0023x_3 + 0.0383x_4 - 0.2868x_5 - 0.0128x_6 + 0.030x_7 + 0.0408x_8 - 0.2083x_9$

Secciones 11-4 y 11-5

11-13. **a)** $f_0 = 260$, se rechaza H_0 **b)** 8.517
c) $se(\hat{\beta}_2) = 0.0861$, $se(\hat{\beta}_1) = 0.195$
d) Se rechaza H_0, ambas son significativas

11-15. **a)** $f_0 = 25.7465$, se rechaza H_0, Valor $P < 0.000001$
b) 3.23 **c)** 0.000714
d) Se rechaza H_0, todas las variables son significativas

11-17. **a)** $f_0 = 53.316$, se rechaza H_0
b) Sólo x_1 es significativa

11-19. **a)** $f_0 = 5.106$, no se rechaza H_0
b) 242.716 **c)** No se rechaza H_0

11-21. **a)** $f_0 = 7.714$, no se rechaza H_0
b) $f_0 = 1.47$, no se rechaza H_0
c) 561.28, 650.14

11-23. **a)** $f_0 = 840.55$, se rechaza H_0
b) 12.1078 **c)** x_2 no es significativa

11-25. **a)** $f_0 = 101.79$, se rechaza H_0
b) "Puntos" es significativa

Secciones 11-6 y 11-7

11-27. **a)** (–0.00657, –0.00122)
b) 0.49765 **c)** (7.16, 9.22)

11-29. **a)** $-0.26718 \le \beta_1 \le 1.47793$;
$-3.61373 \le \beta_2 \le 21.4610$;
$-4.21947 \le \beta_3 \le 7.09438$;
$-1.72211 \le \beta_4 \le 1.74932$
b) (263.78, 311.34)
c) (243.711, 331.413)

11-31. **a)** β_2 : (–0.595, 0.535);
β_3 : (0.229, 0.812);
β_4 : (–0.216, 0.013);
β_5 : (–7.298, 2.977)
b) (7.982, 10.009)
c) (6.8481, 11.143)

11-33. **a)** Temp: (–0.00003, 0.00012);
Soaktime: (0.00203, 0.00288);
Soakpct: (–0.02306, 0.05976);
Difftime: (0.00501, 0.01056);
Diffpct: (–0.01969, 0.01342)
b) (0.026, 0.0234)

11-35. **a)** (0.3882, 0.5998)
b) $\hat{y}' = -5.767703 + 0.496501x_{\text{puntos}}$
c) (0.4648, 0.5282)

Sección 11-8

11-37. **a)** 0.82897
11-39. **a)** 0.86179 **b)** 0.920453
11-41. **b)** 99.37%
c) $\hat{y}' = 6.22489 - 0.16647x_1 - 0.000228x_2 + 0.157312x_3$

11-43. **a)** 95.82% **b)** La normalidad es aceptable
c) La última observación ejerce influencia

11-45. **a)** 0.12 **b)** Los puntos 17 y 18

Secciones 11-9 y 11-10

11-47. **a)** $\hat{y}' = -1.633 + 1.232x - 1.495x^2$
b) $f_0 = 1\ 858\ 613$, se rechaza H_0
c) $f_0 = -601.64$, se rechaza H_0

11-49. **a)** 802.943
b) $\hat{y}' = -26\ 204 + 189.09x - 0.331x^2$
c) Son iguales

11-51. **a)** $\hat{y}' = -1.769 + 0.421x_1 + 0.222x_2 - 0.128x_3 - 0.02x_1x_2 + 0.009x_1x_3 + 0.002x_2x_3 - 0.019\,x_1^2 - 0.007\,x_2^2 + 0.001\,x_3^2$
b) Se rechaza H_0
d) $f_0 = 1.612$, no se rechaza H_0

Sección 11-11

11-55. **a)** Mín. $MS_E = 6.58$, $C_p = 5.88$:
$x_2, x_3, x_4, x_5, x_7, x_8, x_{10}$
Mín. $C_p = 5.02$, $MS_E = 7.97$;
x_5, x_8, x_{10}
b) $\hat{y}' = 34.434 - 0.048x_1$,
$MS_E = 8.81$, $C_p = 5.55$
c) El mismo del inciso b)
d) $\hat{y}' = 0.341 + 2.862x_5 + 0.246x_8 - 0.010x_{10}$,
$MS_E = 7.97$, $C_p = 5.02$

11-57. **a)** $\hat{y}' = 4.656 + 0.511x_3 - 0.124x_4$
b) El mismo del inciso a)
c) El mismo del inciso a)
d) Todos los modelos son iguales

11-59. **a)** $\hat{y} = -0.304 + 0.083x_1 - 0.031x_3 + 0.0004x_2^2$,
$C_p = 4.04$, $MS_E = 0.004$
b) $\hat{y} = -0.256 + 0.783x_1 + 0.022x_2 - 0.042x_3 + 0.0008x_3^2$,
$C_p = 4.66$, $MS_E = 0.004$

11-61. **a)** $\hat{y} = -3.517 + 0.486x_1 - 0.156x_9$,
$C_p = -1.67$
b) $\hat{y} = -5.964 + 0.495x_1 + 0.025x_7 - 0.163x_9$,
$MS_E = 1.67$
c) $R^2 = 0.98448$, sí

11-63. **a)** Sí, $C_p > p$
b) Sí, tanto MS_E como C_p se reducen

Ejercicios complementarios

11-65. **a)** $f_0 = 1\ 323.62$, se rechaza H_0
b) x_4 es significativa

11-67. **a)** $\hat{y}' = -0.908 + 5.482x_1^* + 1.126x_2^* - 3.920x_3^* - 1.143x_4^*$
b) x_1, x_2, x_3, x_4 son todas significativas

11-71. Sí, los VIF de x_3 y x_5 exceden 10

11-73. **a)** Se rechaza H_0
b) No se rechaza H_0 **c)** No

CAPÍTULO 12

Sección 12-2

12-1. **a)** Se rechaza H_0
b) El modelo es satisfactorio

12-3. **a)** Se rechaza H_0 **b)** 0.0005

12-5. **a)** Se rechaza H_0
c) (141.96, 148.05), (7.36, 24.14)

12-7. **a)** No se rechaza H_0 **b)** 0.2138

12-9. **a)** Se rechaza H_0 **b)** 0.0022
d) (71.02, 79.98) **e)** (8.423, 26.327)

Sección 12-3

12-11. 3.32 3.93 4.42

12-13. 2 666.25 2 933.75 2 971.00 3 156.25

12-15. 17.4 20.8 28.2

12-17. 75.5 82.125 92.875

12-19. 129.25 131.5 145 145.25 145.25

12-29. 12.3 13.38 14.8 14.8

Sección 12-4

12-31. **a)** Se rechaza H_0 **b)** 0.0141 **c)** 0.0148
b) Los residuales son aceptables

12-33. **a)** No se rechaza H_0 **b)** 0
c) 23.9994
d) Menor variabilidad de la sustancia química 4

Sección 12-5

12-35. **a)** Se rechaza H_0
c) Cierto indicio de una mayor variabilidad con la boquilla más baja

12-37. **a)** No se rechaza H_0
b) Cierto indicio de variabilidad al incrementar la magnitud de la respuesta

12-39. **a)** No se rechaza H_0
b) La observación 20 es inusual

Sección 12-6

12-41. $n = 5$

12-43. 0.6

12-45. $n = 5$

Ejercicios complementarios

12-47. **a)** Se rechaza H_0 **b)** 0.56 es un punto atípico

12-49. **a)** No se rechaza H_0 **b)** 0.2107
d) 4.037, 0.342

12-51. **a)** Se rechaza H_0
b) –0.19 1.14 1.34 1.74 1.82 2.12

12.53. **a)** 0.3 **b)** $n \cong 11$

CAPÍTULO 13

Sección 13-4

13.1. **a)**

Factor	Temp	Mat	Temp × Mat
f_0	28.968	7.911	3.56

c) Los residuales son satisfactorios

13-3. **a)**

Factor	Cristal	Fósforo	Cristal × Fos.
f_0	273.8	8.842	1.263

13-9. **a)**

Factor	Dopado	Revenido	Dopado × Revenido
f_0	25.231	543.52	3.54

Sección 13-5

13-11. **a)** A–Madera dura B–Tiempo de cocción C–Libertad en la orientación de la fibra

Factor	A	B	C	AB	AC	BC	ABC
f_0	7.553	31.313	19.706	2.889	2.937	0.947	0.940
Valor P	0.0042	0	0	0.0816	0.0495	0.4063	0.4634

Sección 13-6

13-13. **a)** Las variables B, C y AC son significativas

b) $\hat{y} = 413.125 + 9.124x_A + 42.125x_B + 35.875x_C - 59.625_{AC}$

13-17. **b)** Los efectos importantes son A, B, D, E, AB.
c) Deberán incrementarse A, B, D, E.

13-19. **a)** Las variables importantes son A, B y AB.
d) Los residuales del modelo al utilizar A, B y AB son aceptables.

13-21. **a)** $F = 84.64$, valor $P = 0.0008$, se rechaza H_0
b) 8.20

Sección 13-7

13-23. **a)** Bloque 1: (1) ab ac bc;
Bloque 2: a b c abc;
A = 13.75, B = 127.75,
C = 97.75, AB = –21.25,
AC = –137.25, BC = –52.25

13-25. Bloque 1: (1) abc ad bc;
Bloque 2: ac abd cd;
Bloque 3: ab c bd acd;
Bloque 4: a bc d abcd

13-27. ABC, CDE y ABDE

13-29. **a)** Las variables significativas son C, AB, AC
b) Los residuales del modelo al utilizar A, B, C, AB y AC son aceptables.
c) La segunda réplica pudo haber confundido un efecto diferente en los bloques.
d) Confundir ABC en la Réplica 1 y una interacción de dos factores en la Réplica 2.

Sección 13-8

13-31. **a)** Los efectos importantes son:
D + ABCE = 4.545;
A + BCDE = 1.435;
B + ACDE = –1.465;
AB + CDE = 1.15;
AD + BCD
b) Los residuales del modelo al utilizar A, B, D, AB y AD son aceptables.
c) Al reducir los resultados C y E en dos réplicas de un diseño factorial completo en A, B y D.

13-33. **c)** A: –1.525, B: –5.175, C: 2.275, D: –0.675, E: 2.275
b) Se cuenta con AB y AD para usarlas como error, ya que se encontró que no son significativas.

13-35. El diseño reducido consta de dos réplicas de un diseño 2^{4-1} porque ABCE es una palabra en la relación de definición completa.

13-37. A + BD + CE + ABCDE = 1 450;
B + AD + CDE + ABCE = 3 650;
C + AE + BDE + ABCD = –150;
D + AB + BCE + ACDE = 1 750;
E + AC + BCD + ABDE = 650;
BC + DE + ABE + ACD = 250;
BE + CD + ABC + ADE = –50

13-39. A + BD + CE; B + AD + CF;
C + AE + BF; D + AB + EF;
E + AC + DF; F + BC + DE;
AF + BE + CD

Sección 13-9

13-41. **b)** $\hat{y} = 79.947 - 1.115x_1 - 4.208x_2 - 1.801x_1x_2 + 5.707\,x_1^2 - 5.743\,x_2^2$,
f_0 = corrida número 6 es un punto atípico.
c) El rendimiento se incrementa cuando x_1 se decrementa y x_2 está cerca de su nivel cero.

13-43. **a)** $1 - 2.5x_1 < x_2 < 2.25 - 1.5x_1$
b) Sea $x_1 = 1.5$, $x_2 = -1.5$

13-45. **a)** $\hat{y} = 327.57 + 177x_A + 109.425x_B + 131.44x_C + 66x_{AB} + 75.47x_{AC} + 43.58x_{BC}$
b) $\hat{y} = 48 + 11.53x_A - 15.32x_B + 29.19x_C$

Ejercicios complementarios

13-47. **a)**

Factor	Engrane	Tiempo	Engrane × Tiempo
f_0	3.905	25.657	0.260

c) Los residuales son aceptables

13-49. **a)**

Factor	Tope	Alineación	Tope × Alineación
f_0	3.89	10.06	0.875

b) El tope de 75 o 130 con el método 3, para minimizar la respuesta.

13-51. **b)** Los efectos importantes son A = 5.7625, B = 5.5625, C = 6.2375

c) $MS_E = 14.03$

d) $\hat{y} = 72.89 + 2.88x_A + 2.78x_B + 3.19x_C$

e) La variabilidad es mayor cuando A está en el nivel alto.

13-55. Generadores: E = BCD; F = ACD; G = ABC; H = ABD

13-57. **a)** Los efectos importantes son A, C, D, AC y AD

b) $MS_E = 1.625$

c) Los residuales son aceptables

13-59. **a)** D = ABC, los efectos principales son alias de las interacciones de tres factores, con excepción de E que es alias de ABCDE. Otros alias son Ab + Cd, AC + BD, AD + BC cuando se ignoran los efectos de orden superior.

CAPÍTULO 14

Sección 14-2

14-1. No, valor $P = 0.109$

14-3. Sí, valor $P = 0.0002$

14-5. **a)** Valor $P = 0.151$

b) Valor $P = 0.099$ sin corrección de continuidad, 0.151 con corrección de continuidad.

14-7. Sin corrección de continuidad: Valor $P = 0.250$, con corrección de continuidad: 0.359

14-9. Valor $P = 0289$, no se rechaza H_0

14-11. Sin corrección de continuidad: Valor $P = 0.00047$, con corrección de continuidad: 0.0019

14-13. Se rechaza H_0

14-15. Sin corrección de continuidad: Valor $P = 0.0004$, con corrección de continuidad: 0.0019

14-17. **a)** 0.0025 **b)** 0.115

c) La prueba del signo que rechaza H_0 si $R^- \leq 1$ tiene un valor de $\alpha = 0.011$ con base en la distribución binominal.

d) 0.487

14-19. Valor $P = 0.0075$

Sección 14-3

14-21. $w = 80.5$, no se rechaza H_0

14-23. $w = 5$, se rechaza H_0

14-25. $w = 17.5$, se rechaza H_0

14-27. $w = 1$, se rechaza H_0

Sección 14-4

14-29. $w = 38$, no se rechaza H_0

14-31. Valor $P = 0.716$

14-33. $w = 73$, se rechaza H_0

14-35. Valor $P = 0.13$

Sección 14-5

14-37. Se rechaza H_0

14-39. No se rechaza H_0

14-41. Valor $P = 0.018$

Ejercicios complementarios

14-43. $w = 5$, no se rechaza H_0

14-45. Valor $P = 0.133$, no se rechaza H_0

14-47. $r = 1$, no se rechaza H_0
14-49. $w = 4$, se rechaza H_0
14-51. $w = 55$, se rechaza H_0
14-53. Valor $P = 0$, se rechaza H_0
14-55. No se rechaza H_0
14-57. Se rechaza H_0

CAPÍTULO 15

Sección 15-5

15-1. **a)** Carta $\bar{x}$: UCL = 37.5789,
CL = 34.32, LCL = 31.0611,
Carta R UCL = 11.9461,
CL = 5.65, LCL = 0
b) Un punto fuera de los límites de control en la Carta $\bar{x}$ con límites de control revisados:
Carta $\bar{x}$: UCL = 37.4038,
CL = 34.0947, LCL = 30.7857,
Carta R UCL = 12.1297,
CL = 5.73684, LCL = 0

15-3. **a)** Carta $\bar{x}$: UCL = 15.6299,
CL = 10.7125, LCL = 5.79506,
Carta R UCL = 15.4047,
CL = 6.75, LCL = 0
b) Ningún punto está fuera de los límites de control.

15-5. **a)** Carta $\bar{x}$: UCL = 242.78,
CL = 223, LCL = 203.22,
Carta R UCL = 72.51,
CL = 34.286, LCL = 0
b) Media: 223,
desviación estándar: 14.74

15-7. **a)** Carta $\bar{x}$: UCL = 7.44253,
CL = 6.2836, LCL = 5.12467,
Carta R UCL = 2.9153,
CL = 1.1328, LCL = 0
b) Ningún punto está fuera de los límites de control.

Sección 15-6

15-9. **a)** Carta x: UCL = 92.1893,
CL = 81.6667, LCL = 71.144
Carta MR: UCL = 12.9324,
CL = 3.95652, LCL = 0
el proceso está bajo control.
b) Media: 81.6667,
desviación estándar: 3.508

15-11. **a)** Carta x: UCL = 552.12,
CL = 500.6, LCL = 449.088
Carta MR: UCL = 63.308,
CL = 19.3684, LCL = 0
el proceso está bajo control.
b) Media: 500.6,
desviación estándar: 17.17

Sección 15-7

15-13. **a)** PCR = 1.13, $PCR_K = 1.10$
b) 0.075%

15-15. PCR = 0.508, la capacidad del proceso es pobre

15-17. PCR = 0.90, $PCR_K = 0.84$
la capacidad del proceso es marginal

15-19. PCR = 0.50, $PCR_K = 0.36$,
la capacidad del proceso es pobre

15-21. PCR = 0.49, $PCR_K = 0.47$,
la capacidad del proceso es pobre

Sección 15-8

15-23. **a)** Carta U: UCL = 3.81116,
CL = 1.94167, LCL = 0.0721748
b) Hay dos puntos fuera de control.
Los límites de control revisados son
UCL = 3.463, CL = 1.709,
LCL = 0

15-25. **a)** Carta C: UCL = 19.0558,
CL = 9.708, LCL = 0.360874
Hay dos puntos fuera de control como en la carta U.

Sección 15-9

15-27. **a)** 0.218 **b)** 4.59

15-29. **a)** 0.2946 **b)** 1.42

15-31. **a)** 0.2451 **b)** 4.08

15-33. **a)** 0.96327 **b)** 1.04

15-35. **a)** 0.16603 **b)** 6.02

Sección 15-10

15-37. **a)** 0.169548
b) $H = 0.85$, $K = 0.0848$

15-39. **a)** 38 **b)** 10.4

Ejercicios complementarios

15-41. **a)** Carta $\bar{x}$: UCL = 64.0179, CL = 64.0002, LCL = 63.9862
Carta R: UCL = 0.0443678, CL = 0.01724, LCL = 0
b) Media: 64.0002, desviación estándar: 0.0102
c) PCR = 1.63 **d)** $PCR_K = 1.63$
e) 0.025 **f)** 12.18

15-43. **a)** UCL = 4.81991, CL = 2.64, LCL = 0.460092
b) El proceso está fuera de control. Los límites de control revisados son UCL = 4.54206, CL = 2.44444, LCL = 0.346827
c) UCL = 4.18143, CL = 2.64, LCL = 1.09857
El proceso está fuera de control. Los límites de control revisados son UCL = 3.51897, CL = 2.13333, LCL = 0.747693

15-45. **a)** 0.125 **b)** 0.004

15-49. **a)** 0.10 **b)** 0.09 **c)** 0.2133

15-51. **a)** 0.085 **b)** 0.444

15-55. 0.00135

15-57. $K = 8.5853$, $H = 85.85$

Apéndice B-I

I-1. 120

I-3. 144

I-5. 14 400

I-7. **a)** 416965528 **b)** 113588800

I-9. **a)** 21 **b)** 2 520 **c)** 720

I-11. **a)** 1 000 **b)** 160 **c)** 720

I-13. **a)** 0.461

I-15. **a)** 0.0082 **b)** 0.0082

Apéndice B-II

II-1. **b)** $(m + 1)/2$, $(9m^2 + 10m - 13)/12$

II-3. **b)** $(1 - p)/p^2$

II-5. $E(X) = 1/4$ $V(X) = 5/16$

II-7. **b)** $E(X) = 2/\lambda^2$, $V(X) = 1/\lambda^2$

II-9. **a)** $(1 - t/\lambda)^{-n}$ **b)** Gamma

Apéndice B-III

III-1. 1/4 en $y = 3, 5, 7, 9$

III-3. **a)** $(y - 10)/72$ **b)** 18

III-5. **b)** 1/4

III-7. **a)** $b^2/2$

III-11. **a)** $a/4y$ **b)** $-a/4 \ ln(a/4)$

Apéndice B-V

V-1. **a)** Normal con media $\dfrac{(\sigma^2\mu_0 + \sigma_0^2 x)}{\sigma_0^2 + \sigma^2}$
y varianza $\dfrac{(\sigma^2\sigma_0^2)}{\sigma_0^2 + \sigma^2}$
b) $\dfrac{(\sigma^2\mu_0 + \sigma_0^2 x)}{\sigma_0^2 + \sigma^2}$

V-3. Gamma con parámetros $m + x + 1$ y $1 + (m + 1)/\lambda_0$

V-5. **a)** 5.046 **b)** (4.87, 5.22)

Apéndice B-VII

VII-1. **b)** $\hat{\sigma} = 0.082$ $\hat{\sigma}^2_{\tau\beta} = 0.013$
$\hat{\sigma}^2_{\beta} = 0.532$ $\hat{\sigma}^2_{\tau} = 0.362$

VII-3. **d)** $\hat{\sigma}^2 = 0.0596$ $\hat{\sigma}^2_{\tau\beta} = 0.0077$
$\hat{\sigma}^2_{\beta} = -0.0137 \rightarrow 0$

Índice

Abusos comunes de la regresión, 446
Adecuación del modelo, 461
Alias, 686
Análisis de regresión:
  adecuación del modelo, 461, 521
  cuadrados medios ponderados, 482
  estimaciones de mínimos cuadrados, 434, 488
  prueba de falta de ajuste, 465
  regresión lineal múltiple:
    enfoque matricial, 491
    estimación de parámetros, 488
    intervalos de confianza en la, 515-518
    intervalos de predicción, 518
    modelos polinomiales, 530
    prueba de hipótesis en la, 506-513
    punto atípico, 462
    regresión y correlación, 472
  regresión lineal simple:
    estimación de la pendiente y la ordenada al origen, 435
    estimación en intervalos de confianza para la pendiente y la ordenada al origen, 454
    intervalo de confianza para la pendiente de la recta de regresión, 455
    intervalo de predicción, 458
    prueba de hipótesis en la, 447-452
  selección de variables, 539-549
  transformación a una línea recta, 471
  variables indicadoras, 533
Análisis de residuales, 461, 522, 576, 652
Análisis de varianza, 450, 507, 562, 628
  comparación gráfica de medias, 581
  componentes de la varianza, 589, B-36, B-40
  contrastes ortogonales, 583, 584
  diseño de bloques aleatorizados, 595
  experimento con un solo factor (una dirección), 563:
    modelo de efectos aleatorios, 566, 589
    modelo de efectos fijos, 566
  experimentos con factores múltiples (multidirecciones), 641
  experimentos factoriales con dos factores (en dos direcciones):
    modelo con efectos fijos, 627
    modelo de efectos aleatorios, 638, B-35
    modelo mixto, 638, B-39
  experimentos factoriales con tres factores, 641
  método de la diferencia significativa mínima (LSD) de Fisher, 586, 603, 635
  métodos no paramétricos, 747
  modelo de efectos aleatorios, 589, B-35
  modelo de regresión, 432, 507
  potencia de la prueba para el modelo con efectos aleatorios, 610
  verificación de la adecuación del modelo, 576, 604, 635
Aproximación binomial de la distribución hipergeométrica, 143
Aproximación de Poisson a la distribución binomial, 135
Ascenso más pronunciado, 707
Asociación lineal, 239
Axiomas de probabilidad, 66

Bigote, 37
Bloque principal, 679
Bootstrap (cargador inicial), 270, 325

Carta *C* para defectos, 793
Carta de control de suma acumulada (CUSUM), 797
Carta de control para mediciones individuales, 776
Carta de tolerancia, 781
Carta *P* o carta de la proporción defectuosa, 788
Carta *R*, 768, 770
Carta *U* (carta de control de defectos por unidad), 790
Carta *X*, 768, 769
Cartas de control de atributos, 761, 788
Cartas de control de Shewhart, 759
Cartas de control de variables, 761
Catorce puntos de Deming, 810
Causa común, 758

Causa especial, 758
Causa y efecto, 382
Causas asignables, 758
Causas fortuitas, 757
Codificación de los datos, 25
Coeficiente de confianza, 289
Coeficiente de correlación, 240, 473
  interpretación de, 240-244
Coeficiente de correlación múltiple, 522
Coeficiente de determinación, *ver* R^2
Coeficiente de regresión, 432
Colinealidad múltiple, 554
Combinación lineal, 249
Comparación gráfica de medias, 581
Comparación múltiple, 581
Componentes de la varianza, 589, 591, B-36, B-40
Confusión o mezclado, 677
Conjunto, 52
  complemento de un, 52
  intersección de un, 52
  unión de un, 52
Construcción del modelo en regresiones, 539
Contraste de definición, 678
Contrastes, 583
Control estadístico de procesos, 23, 756, 757
Corrección para poblaciones finitas, 132
Correlación y regresión, 472
Corrida, 765
Corrupción de un proceso, 21
Covarianza, 237, 238, 504
C_p, 541
Cuadrados medios, 451
Cuadrados medios esperados:
  diseño de bloques completos aleatorizados, 597, 599, 605
  experimento con un solo factor (una dirección), 563:
    modelo de efectos aleatorios, 566
    modelo de efectos fijos, 566
  experimentos con dos factores (dos direcciones):
    modelo con efectos fijos, 627
    modelo de efectos aleatorios, B-35
    modelo mixto, B-39
    modelo restringido, B-39
Cuadrados medios ponderados, 481
Cuartiles, 30
Curva de operación característica (OC), 308, 316, 335, 344, 379, 394, 414, 609, 611
  para el análisis de varianza del modelo con efectos aleatorios, A-25
  para el análisis de varianza del modelo con efectos fijos, A-21
  para probar:
    la igualdad de las medias de dos distribuciones normales, σ_1 y σ_2 conocidas, A-12
    la igualdad de las medias de dos distribuciones normales, σ_1 y σ_2 desconocidas, A-14
    la igualdad de las varianzas de dos distribuciones normales, A-19
    la media de una distribución normal, σ_2 conocida, A-12
    la media de una distribución normal, σ_2 desconocida, A-14
    la varianza de una distribución normal, A-16
CUSUM tabular (suma acumulada), 800

Datos con multivariables, 36
Datos unimodales, 35
Diagrama de causa y efecto, 757, 807
Diagrama de concentración de defectos, 757, 807
Diagrama de dispersión, 431
Diagrama de Pareto, 36, 757, 806
Diagrama de puntos, 3
Diagrama de tallo y hoja, 27, 30
Diagrama o carta de control, 22, 758
  Carta *C*, 793
  Carta de suma acumulada (CUSUM), 797
  Carta del promedio móvil ponderado exponencialmente (EWMA), 816
  Carta *NP*, 790, 816
  Carta *P*, 788
  Carta *R*, 768, 770
  Carta *U*, 790
  Carta *X*, 768, 769
  estandarizada, 817
  mediciones individuales, 776
  rango móvil, 776
Diagrama punto-dígito (*digidot*), 41
Diagramas de árbol, 56
Diagramas de Venn, 55
Desigualdad de Chebyshev, 254
Desviación estándar, 5, 7, 108, 112, 153
Diseño central compuesto, 711, 713-715
Diseño de bloques, 595, 677
Diseño de bloques completos aleatorizados, 595
Diseño de resolución, 693
Diseño desbalanceado, 573
Diseño factorial, 17, 618, 621, 647
Diseño factorial 2^k, 647
Diseño factorial fraccionado, 684
Diseño girable, 715
Diseño ortogonal, 660
Diseños con una sola réplica, 636, 666
Dispersión o variabilidad, 3
Distribución binomial, 112, 115
  aproximación normal a la, 174
  media, 118
  propiedades de, 115-119
  varianza, 118
Distribución binomial negativa, 124
  media, 126
  varianza, 126
Distribución continua uniforme, 155
  media, 155
  varianza, 155
Distribución de dos variables o bivariable, 200, 245
Distribución de Erlang, 188
Distribución de frecuencia, 32
Distribución de frecuencia relativa, 33
Distribución de Gauss o gaussiana, 157

Distribución de muestreo, 261, 281, 282, 309
Distribución de Poisson, 135
aproximación normal, 174
media, 139
propiedad reproductiva de la, B-13
varianza, 139
y la distribución binomial, 135
Distribución de probabilidad de una variable aleatoria, 99, 145
Distribución de Weibull, 193
aplicación de la, 193
media, 194
varianza, 194
Distribución discreta uniforme, 110
Distribución exponencial, 179, 180
propiedades de, 182-184
Distribución *F*, 410, B-26
media de la, 411
tabla de la, A-7
varianza de la, 411
Distribución gamma, 189
media, 191
varianza, 191
y distribución ji-cuadrada, 191
Distribución geométrica, 121, 122
media, 123
propiedades de, 121-124
varianza, 123
Distribución hipergeométrica, 129
media, 131
varianza, 131
y distribución binomial, 132
Distribución ji-cuadrada, 191, 341
media de la, 341
prueba de la bondad del ajuste, 356
tabla de, A-5
varianza de la, 341
y distribución gamma, 191
Distribución multinomial, 215
Distribución normal, 157
aproximación normal a la distribución binomial, 174
aproximación normal a la distribución de Poisson, 174
combinación lineal de, 249
de dos variables, 245-248
distribución normal estándar, 160, 163, 164
media, 159
propiedad reproductiva de la, 252
tabla de, A-3
varianza, 159
Distribución normal de dos variables, 245-248
Distribución normal estándar, 160
Distribución normal estándar acumulada, 161
Distribución posterior, B-28, B-29
Distribución previa, B-28, B-29
Distribución *t*, 330, B-24
media, 330
tabla de la, A-6
varianza, 330
Distribución *t* no central, 335
Distribución uniforme:
continua, 155
discreta, 110
Distribuciones:
binomial, 115, A-2
binomial negativa, 124, 125, A-2
continua uniforme, 155, A-2
de Erlang, 187, 188, A-2
de Poisson, 135, 136, A-2
de Weibull, 193, A-2
discreta uniforme, 110, A-2
exponencial, 179, 180, A-2
F, 410
gamma, 187, 189, 190, A-2
geométrica, 122, A-2
hipergeométrica, 129, A-2
ji-cuadrada, 191, 341
multinomial, 214, 215
normal, 157, A-2
normal de dos variables, 245-248
t, 330
Distribuciones de probabilidad conjunta, 199, 200, 211, 218, 232
Distribuciones de probabilidad marginal, 202, 222, 234, 247, B-20

Ecuaciones normales, 435, 489-493
Efecto principal, 622
Efectos de interacción, 18, 485, 622, 624
Eficiencia relativa de un estimador, 272
Elemento identidad, 660
Eliminación hacia atrás, 548
Enfoque bayesiano, B-28
Ensayo de Bernoulli, 113
Error cuadrado medio de un estimador, 271
Error estándar, 268, 446, 505, 665
Error estándar de un diseño 2^k, 665
Error estándar estimado, 268, 505
Errores tipo I y tipo II, 297, 298
Espacio muestral, 52
discreto, 62
Estadística, campo de la, 2
Estadístico, 263
Estadístico de orden, 292
Estadístico de prueba, 309
Estimación de parámetros, 260
Estimación de parámetros:
estimador de Bayes, B-29
estimador de máxima verosimilitud, 274
estimador insesgado, 265
estimador insesgado de varianza mínima, 267
mínimos cuadrados, 444
Estimación del rango móvil de σ, 776
Estimación puntual, 261, 264
Estimador consistente, 292
Estimador de Bayes, B-29
Estimador de máxima verosimilitud, 274
de parámetros de distribuciones:
de Bernoulli, 274
exponencial, 276
normal, 275, 277
uniforme, 280
función de verosimilitud, 274
propiedades del, 278
Estimador insesgado, 265, 309, 444, 504
Estimador insesgado de varianza mínima, 267
Estimador óptimo, 272
Estimador puntual, 264
Estimadores insesgados, 265, 267
Estrategia de experimentación, 562
Estudio analítico, 10
Estudio enumerativo, 10
Estudio observacional, 11, 382
Estudios de la capacidad de un proceso, 761, 781

Experimento, 11, 17, 49
Experimento aleatorio, 50, 52
Experimento comparativo, 295
Experimento completamente aleatorizado, *ver* Experimento aleatorizado
Experimento del embudo de Deming, 21
Experimento diseñado, 11, 17, 562, 618, 757, 807
Experimento factorial con dos factores, 622
Experimento factorial con tres factores, 641
Eventos, 54
  complemento de, 54
  independientes, 84, 85, 86
  intersección de, 54
  mutuamente excluyentes, 58
  unión de, 54
Eventos independientes, 84
Eventos mutuamente excluyentes, 58, 71
EWMA, carta de control del promedio móvil ponderado exponencialmente, 816
Experimentación secuencial, 563, 687, 707
Experimento aleatorizado, 382, 563, 564, 627
Experimento de caracterización de un proceso, 618, 619
Experimento de optimización, 620, 705
Experimento exhaustivo de depuración o tamiz, 563, 619
Experimento factorial con interacción entre los factores, 622
Experimento factorial fraccionado, 19, 684

Factor de inflación de la varianza, 554, 555
Factorial fraccionado saturado, 703
Falta de ajuste, 465
Fórmulas para el tamaño de una muestra:
  para intervalos de confianza, 324, 385
  para probar hipótesis, 315, 380, 381
  para proporciones, 349, 353
Fracción alternativa, 687
Fracción principal, 687
Frecuencia relativa, 62
Fuentes de variabilidad, 3
Función de densidad, 155
Función de densidad de probabilidad, 145, 146
  conjunta, 218, 232
  marginal, 222, 234
Función de densidad de probabilidad condicional, 224
Función de densidad de probabilidad conjunta, 218, 219, 232
Función de densidad de probabilidad marginal, 222, 234
Función de distribución, *ver* Función de distribución acumulada
Función de distribución acumulada, 104, 149
  gráfica de, 105, 150
  propiedades de la, 104
  y función de densidad de probabilidad, 151
Función de masa, 100, 104
Función de masa de probabilidad, 100
Función de masa de probabilidad condicional, 205
Función de masa de probabilidad conjunta, 201, 211-212
Función de masa de probabilidad marginal, 203, 212
Función de potencia, 308
Función de verosimilitud, 274
Función gamma, 190
Función generadora de momentos, B-7, B-8
  para la distribución de la suma de dos o más variables aleatorias independientes, B-12
  propiedad de unicidad de la, B-11
Funciones de variables aleatorias, B-15
  variable aleatoria continua, B-18, B-20
  variabale aleatoria discreta, B-16, B-17
Generador del diseño, 686, 693, 698
Grados de libertad, 8
Gráfica de contorno, 485
Gráfica de frecuencias acumuladas, 34
Gráfica de probabilidad normal, 170, 171, 333, 577, 654, 669, 671-682, 690, 691
Gráfica de probabilidades, 170, 356; *ver también* Gráfica de probabilidad normal
Gráfica de series de tiempo, 40
Gráficas de caja, 37
Gráficas de interacción de dos factores, 624

Hipótesis:
  alternativa, 295
  nula, 295
  procedimiento general de las pruebas de, 307, 318
  prueba de, 290, 295, 296
Hipótesis alternativa, 295
Hipótesis de dos colas, 295-296, 304
Hipótesis de una cola, 304
Hipótesis nula, 295
Histograma, 32, 33, 146, 781

ICP (índice de la capacidad del proceso), 783
ICP_k (índice de la capacidad potencial del proceso), 785
Índice de la capacidad del proceso (ICP), 783, 785
Inferencia estadística, 10
Interacción generalizada, 682
Interpretación física de la media muestral, 4
Intervalo de Bayes, B-31
Intervalo de confianza, 289, 320, 324
  de dos colas, 320
  de una cola, 320
  para el cociente de las varianzas de dos distribuciones normales, 415
  para la diferencia:
    de dos proporciones, 421
    en dos medias, varianzas conocidas, 383

en las medias de dos distribuciones normales, varianzas desconocidas pero iguales, 395
en las medias de dos distribuciones normales, varianzas desconocidas y diferentes, 397
para la media:
de una distribución normal, varianza desconocida, 336
varianza conocida, 321
para la pendiente y la ordenada al origen en una regresión lineal simple, 454, 455
para la recta de regresión lineal simple, 455
para la varianza de una distribución normal, 345
para medias de tratamientos en un análisis de varianza, 573-575
para observaciones pareadas, 406
para una proporción, 347
relación entre las pruebas de hipótesis, 322
técnica *bootstrap* (cargador inicial), 325

Intervalo de confianza superior e inferior, 324
Intervalo de predicción, 372, 458
Intervalo de tolerancia, 373
Intervalos de clase, 32
Invarianza del estimador de verosimilitud máxima, 279

Jacobiano de una transformación, B-18

Límites de advertencia en cartas de control, 767
Límites de control, 22, 758
Línea central (LC) en un diagrama (o carta) de control, 22, 758
Localización, 3
Longitud de la corrida promedio, 794, 803
Máscara V, 799
Matriz de covarianza, 504
Matriz "gorro" en regresiones, 525
Media condicional, 207, 227
Media fracción de un experimento factorial, 20, 684
Media muestral, 3, 4
Media o valor esperado, 108
Media poblacional, 4, 5, 106
Media recortada, 48
Mediana muestral, 30
Medida de la distancia de Cook, 527
Método científico, 1
Método de la ingeniería para resolver problemas, 1
Método de mínimos cuadrados, 15, 434, 488
Método de suma extra de los cuadrados, 510
Métodos de distribución libre, 727
Métodos paramétricos, 726
Mínimos cuadrados, *ver* Método de mínimos cuadrados
Moda muestral, 30
Modelo, 9, 50
Modelo de probabilidad, 9
Modelo de regresión, 15, 432, 484, 652
Modelo de regresión lineal simple, 432
Modelo de regresión múltiple, 484
Modelo de superficie de respuesta de primer orden, 706
Modelo de superficie de respuesta de segundo orden, 706, 711
Modelo empírico, 13, 14, 432
Modelo lineal probabilístico, 432
Modelo mecanicista, 13
Modelos de regresión polinomiales, 530
Momento de una variable aleatoria, B-8
Muestra, 4, 260, 261, 262
cuartiles, 30
desviación estándar muestral, 5
media muestral, 3, 4
mediana muestral, 30
moda muestral, 30
rango muestral, 8
valor más grande, 8
valor más pequeño, 8
varianza muestral, 5
Muestra aleatoria, 261, 262
generación de muestras aleatorias, 262, 263
Muestreo, 261
Muestreo con y sin reemplazo, 54, B-5

Nivel de significación o amplitud, 298
Niveles del factor, 562
Notación geométrica del diseño factorial, 648

Observaciones influyentes, 526
Operador varianza, 108, 153

Papel probabilidad, 170
Pensamiento estadístico, 3, 9
Percentil, 31
Pérdida de calidad, 197
Permutación, B-2
Población, 4, 260, 262
Potencia de una prueba estadística, 304, 308
Precisión del estimador, 290
Predictor, 434
Principio de efectos esparcidos, 667
Principio del cociente de verosimilitud, B-32
Probabilidad, 9, 49
axiomas de, 66
condicional, 74, 75
regla de probabilidad total, 79, 80
reglas de adición, 68
reglas de multiplicación, 79
Probabilidad condicional, 74, 75, 205
Probabilidad total, 79, 80, 81
Probabilidades subjetivas, B-28
Procedimiento general de las pruebas de hipótesis, 307
Proceso de Poisson, 136
supuestos del, 136
Proceso seis-sigma, 197, 786
Propiedad reproductiva:
de la distribución de Poisson, B-13
de la distribución normal, 252
Propiedad de falta de memoria, 124, 182, 198

Propiedades asintóticas del estimador de máxima verosimilitud, 278
Propiedades de los estimadores, 265
Proyección de diseños 2^k, 666, 691
Prueba de curvatura utilizando puntos centrales, 672
Prueba de la bondad del ajuste, 356
Prueba de la diferencia de significación mínima de Fisher, *ver* Análisis de varianza
Prueba de la suma de rangos de Wilcoxon, 743
Prueba de hipótesis:
  conclusiones robusta y débil, 304
  en la regresión lineal múltiple, 506-513
  en la regresión lineal simple, 447-452
  observaciones pareadas, 402
  relativas a la bondad del ajuste, 466
  relativas a la mediana, 728
  relativas a medias:
    de distribuciones, varianza conocida, 309
    de distribuciones normales, varianza desconocida, 328
    de dos distribuciones, varianzas conocidas, 377
    de dos distribuciones normales, varianzas desconocidas pero iguales, 388
    de dos distribuciones normales, varianzas desconocidas y diferentes, 392
  relativas a varianzas de poblaciones normales, 340, 410
  sobre proporciones, 347, 417
Prueba de Kruskal-Wallis, 747
Prueba de Mann-Whitney, *ver* Prueba de la suma de rangos de Wilcoxon
Prueba de rangos con signo de Wilcoxon, 738
Prueba de signos, 727
Prueba estadística de hipótesis, 12, 295; *ver también* Prueba de hipótesis
Prueba F parcial, 512
Prueba general de la significación de la regresión, 510
Prueba t combinada, 390
Prueba t pareada, 402, 403
Pruebas de falta de ajuste en regresiones, 465
Pruebas de hipótesis e intervalos de confianza, 290, 295, 322
Punto de brazo de palanca en regresiones, 530
Puntos atípicos, 37
Puntos centrales en un diseño 2^k, 671, 672

R^2, 464, 521
R^2 ajustado, 541
Rango intercuartílico, 31
Rango muestral, 8
Rangos, 727, 738, 743, 748, 750
Región crítica, 297, 310
Región de aceptación, 297, 310
Región de rechazo de una prueba, 310
Regla de adición de probabilidad, 68
Regla de multiplicación para la probabilidad, 79
Reglas de zona para cartas de control, 768
Reglas Western Electric, 767
Regresión por pasos, 545
Regresión sobre ascensos (cordilleras), 556
Relación de definición, 686, 694
Réplicas, 17, 564, 627
Representación ordenada de tallo y hoja, 29
Residual studentizado, 470, 524
Residuales, 436, 496, 522
Residuales estandarizados, 462, 522
Resolución del diseño, 693
Resultados igualmente factibles, 63

Selección hacia adelante, 547
Sensibilidad de una prueba estadística, 304
Sesgo, 265
Sesgo o asimetría, 35
Significación de una regresión, 448, 506
Significación estadística y significación práctica, 318
Significación práctica, 318
Sobrecontrol de un proceso, 21
Subgrupo racional, 763
Suma de cuadrados atribuible al error puro, 466
Suma de los cuadrados de los errores, 444
Superficie de respuesta, 705, 713

Tabla de contingencia $r \times c$, 361
  para poblaciones homogéneas, 364
Técnicas de conteo, B-1
Tendencia, 40
Teorema de Bayes, 89, 90
Teorema del límite central, 157, 283
Todas las regresiones posibles, 539
Transformación de rangos, 750
Transformaciones, 47, 471
Tratamientos, 382, 564
Trayectoria del ascenso más pronunciado, *ver* Ascenso más pronunciado

Unimodales, 35

Valor esperado de una variable aleatoria, 108, 153
  media condicional, 227
Valor P, 312, 334
Valor z, 164
Valores críticos, 297
Variabilidad, 3, 23
Variabilidad cíclica, 41
Variable aleatoria, 91, 92
  continua, 92
  discreta, 92, 97
  independiente, 228-229, 250-251
Variable aleatoria normal estándar, 164
Variable de regresión, 432, 434, 485
Variable de respuesta, 434, 485
Variables aleatorias continuas, 144
Variables aleatorias discretas, 97

Variables aleatorias independientes, 208-210, 228-229, 251
criterio para, 209, 235, 242
Variables cualitativas en un modelo de regresión, 533
Variables indicadoras en regresiones, 533
Varianza condicional, 207, 227
Varianza de una variable aleatoria, 106, 107, 153
Varianza muestral, 5
Varianza poblacional, 7, 8
evaluación de, 8